Physical Mathematics II

An Undergraduate Course in Fourier Analysis, Complex Variables, Linear Algebra, and the Calculus of Variations

Weinan Wang
Department of Mathematics, University of Oklahoma
ww@ou.edu

February 15, 2026

Contents

Preface

These notes accompany a one-semester undergraduate course in the mathematical methods of the physical sciences—the second course in a two-semester sequence. The first course (Physical Mathematics I) develops the calculus of vectors and fields, ordinary differential equations, and an introduction to series and special functions. These notes take up where that course leaves off and treat four large subjects that recur throughout physics and engineering: *Fourier analysis*, the *theory of functions of a complex variable*, the *linear algebra of vector and function spaces*, and the *calculus of variations*. A recurring theme ties them together: again and again we are handed a differential equation or an optimization problem coming from a physical system, and the right change of representation—into frequencies, into the complex plane, into an eigenbasis, or into a variational principle—turns something intractable into something we can solve by hand.

Who this is for. The notes assume the standard prerequisites for the course: multivariable calculus through line and surface integrals, and a first exposure to ordinary differential equations and infinite series at the level of Physical Mathematics I. No prior course in complex analysis, abstract linear algebra, or the calculus of variations is assumed; we build each subject from its definitions. A reader who has seen some of this material before will find the early sections of each chapter to be review and can move quickly to the applications.

The text. The course follows D. A. McQuarrie, *Mathematical Methods for Scientists and Engineers* (University Science Books, 2003), drawing on parts of Chapters 4–10 and 17–20. These notes are not a substitute for that book; they are a companion that fixes notation, supplies the derivations in the order we present them in lecture, and collects exercises. Section numbers in the text are occasionally cited so that a reader can locate the corresponding fuller treatment in McQuarrie.

How to read these notes. Sections and results marked with a star ($\star$) are more demanding—they contain a proof or a subtlety that can be skipped on a first reading without loss of continuity, and are included for students who want the underlying justification. Each chapter ends with exercises; some are routine practice with the methods, others ask for a short derivation or a physical interpretation. Solution sketches for selected exercises are collected at the end of the notes. Worked examples in the body of the text are an integral part of the exposition and should be read with pencil in hand.

Plan of the notes. The material is organized into four parts. *Part I (Chapters 1–3)* develops Fourier series and the Fourier transform and applies them to the classical partial differential equations of mathematical physics. *Part II (Chapters 4–9)* is a self-contained introduction to complex analysis: the algebra and geometry of complex numbers, analytic functions and the Cauchy–Riemann equations, contour integration and the Cauchy theory, Laurent series and residues, the evaluation of real integrals, and conformal mapping with its applications to potential theory and fluid flow. *Part III (Chapters 10–13)* treats the linear algebra of finite- and infinite-dimensional vector spaces—inner products, orthogonal functions, eigenvalue problems and the spectral theorem, the matrix exponential, and the description of space in curvilinear coordinates and the language of tensors. *Part IV (Chapters 14–15)* is an introduction to the calculus of variations, from the Euler–Lagrange equation and Lagrangian mechanics to constrained problems and the method of Lagrange multipliers.

A note on conventions. Throughout, $\mathbb{R}$ and $\mathbb{C}$ denote the real and complex numbers; i denotes the imaginary unit and e the base of the natural logarithm. Vectors and matrices are set in boldface. We write $\langle \cdot, \cdot \rangle$ for an inner product and $\|\cdot\|$ for the associated norm. A function is called *piecewise smooth* on an interval if the interval can be cut into finitely many pieces on each of which the function and its first derivative are continuous, with finite one-sided limits at the cut points.

Chapter 1

Fourier Series

1.1 The basic question

A great many physical systems are *periodic*: the displacement of a plucked string, the voltage in an alternating-current circuit, the temperature at a fixed depth in the ground over the course of a year. A function f is *periodic with period* $T > 0$ if

$$f(x+T) = f(x) \qquad \text{for all } x.$$

The smallest such T is the *fundamental period*. The basic question of this chapter is whether an arbitrary periodic function can be built out of the simplest periodic functions we know—sines and cosines—and if so, how.

The answer, discovered by Joseph Fourier in his 1822 study of heat conduction, is a resounding yes, and it is hard to overstate how consequential it has been. The idea is that a sufficiently well-behaved periodic function is an infinite superposition of pure oscillations whose frequencies are integer multiples of one fundamental frequency. Writing a function in this form—resolving it into its *frequency content*—is the single most useful change of representation in all of applied mathematics, and it is the starting point for everything in Part I.

We work on an interval of length $2L$, say $[-L, L]$, and extend periodically to all of $\mathbb{R}$ with period $2L$. The functions

$$1, \quad \cos\frac{\pi x}{L},\ \sin\frac{\pi x}{L}, \quad \cos\frac{2\pi x}{L},\ \sin\frac{2\pi x}{L}, \quad \ldots$$

all have period $2L$ (the n-th pair completes n oscillations on the interval), so any finite or infinite combination of them is again $2L$-periodic. Fourier's claim is that, conversely, essentially every $2L$-periodic function is such a combination.

1.2 The trigonometric Fourier series

Definition 1.1. Let f be a $2L$-periodic function that is integrable on $[-L, L]$. Its *Fourier series* is the trigonometric series

$$f(x) \;\sim\; \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L} \right), \tag{1.1}$$

where the *Fourier coefficients* are

$$a_n = \frac{1}{L}\int_{-L}^{L} f(x)\cos\frac{n\pi x}{L}\,dx \quad (n \geq 0), \qquad b_n = \frac{1}{L}\int_{-L}^{L} f(x)\sin\frac{n\pi x}{L}\,dx \quad (n \geq 1). \tag{1.2}$$

The symbol "$\sim$" is deliberate: at this stage (1.1) only asserts that the coefficients are *defined* by (1.2). Whether the series actually *converges* to f, and in what sense, is a separate question taken up in Section 1.7. The constant term is written $a_0/2$ rather than a_0 so that the single formula for a_n in (1.2) is valid for $n = 0$ as well; with this convention $a_0/2$ is precisely the average value of f over one period.

The coefficient formulas are not pulled from nowhere. They are forced on us by a remarkable property of the sines and cosines, to which we now turn.

1.3 Orthogonality and the coefficient formulas

The reason the coefficients can be computed one at a time—without solving an infinite system of equations—is that the trigonometric functions are *mutually orthogonal* on $[-L, L]$ in the sense of the following lemma. Here and throughout, two functions are called orthogonal on an interval if the integral of their product over that interval vanishes.

Lemma 1.2 (Orthogonality relations)**.** *For integers $m, n \geq 1$,*

$$\int_{-L}^{L} \cos\frac{m\pi x}{L}\cos\frac{n\pi x}{L}\,dx = L\,\delta_{mn}, \qquad \int_{-L}^{L} \sin\frac{m\pi x}{L}\sin\frac{n\pi x}{L}\,dx = L\,\delta_{mn},$$

and for all integers $m \geq 0$, $n \geq 1$,

$$\int_{-L}^{L} \cos\frac{m\pi x}{L}\sin\frac{n\pi x}{L}\,dx = 0,$$

where δ_{mn} is the Kronecker delta (1 if $m = n$, else 0). Moreover $\int_{-L}^{L} 1\,dx = 2L$.

Proof. Use the product-to-sum identities, for example $\cos A \cos B = \frac{1}{2}[\cos(A - B) + \cos(A + B)]$. With $A = m\pi x/L$ and $B = n\pi x/L$,

$$\int_{-L}^{L} \cos\frac{m\pi x}{L}\cos\frac{n\pi x}{L}\,dx = \frac{1}{2}\int_{-L}^{L} \cos\frac{(m-n)\pi x}{L}\,dx + \frac{1}{2}\int_{-L}^{L} \cos\frac{(m+n)\pi x}{L}\,dx.$$

If $m \neq n$, both integrands are cosines of a nonzero integer multiple of $\pi x/L$ and integrate to zero over a full set of periods. If $m = n \geq 1$, the first integral is $\frac{1}{2}\int_{-L}^{L} 1\,dx = L$ and the second still vanishes. The sine–sine relation is identical with the identity $\sin A \sin B = \frac{1}{2}[\cos(A - B) - \cos(A + B)]$, and the sine–cosine integral vanishes because its integrand is an odd function on the symmetric interval $[-L, L]$. □

Granting for the moment that f equals its series (1.1), multiply both sides by $\cos(m\pi x/L)$ and integrate over $[-L, L]$. If we may integrate term by term, Lemma 1.2 annihilates every term on the right except the one containing $\cos(m\pi x/L)$, leaving

$$\int_{-L}^{L} f(x)\cos\frac{m\pi x}{L}\,dx = a_m \cdot L,$$

which is exactly the formula (1.2) for a_m. Multiplying instead by $\sin(m\pi x/L)$ isolates b_m. An orthogonal family lets one read off each coefficient of an expansion by a single inner product—the same mechanism that governs orthogonal polynomials and the eigenfunctions of symmetric operators.

1.4 The complex (exponential) form

Sines and cosines are real and intuitive, but for computation the complex exponential is almost always cleaner, because differentiation and multiplication of exponentials are so simple. Using Euler's formula $\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta$, the pair $\{\cos(n\pi x/L), \sin(n\pi x/L)\}$ can be repackaged into the single family $\{\mathrm{e}^{\mathrm{i}n\pi x/L}\}_{n\in\mathbb{Z}}$.

Definition 1.3. The *complex Fourier series* of a $2L$-periodic integrable function f is

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n\, \mathrm{e}^{\mathrm{i}n\pi x/L}, \qquad c_n = \frac{1}{2L}\int_{-L}^{L} f(x)\, \mathrm{e}^{-\mathrm{i}n\pi x/L}\, dx. \tag{1.3}$$

The coefficient formula again follows from an orthogonality relation, this time the clean statement

$$\int_{-L}^{L} \mathrm{e}^{\mathrm{i}m\pi x/L}\, \overline{\mathrm{e}^{\mathrm{i}n\pi x/L}}\, dx = \int_{-L}^{L} \mathrm{e}^{\mathrm{i}(m-n)\pi x/L}\, dx = 2L\,\delta_{mn}. \tag{1.4}$$

The two forms carry the same information. Comparing (1.1) and (1.3) and matching terms gives the dictionary

$$c_0 = \frac{a_0}{2}, \qquad c_n = \frac{a_n - \mathrm{i}b_n}{2}, \qquad c_{-n} = \frac{a_n + \mathrm{i}b_n}{2} = \overline{c_n} \quad (n \ge 1). \tag{1.5}$$

In particular, when f is real-valued the negative-index coefficients are the complex conjugates of the positive-index ones; the single relation $c_{-n} = \overline{c_n}$ encodes the reality of f.

1.5 Even and odd functions; sine and cosine series

Symmetry cuts the work in half. Recall that f is *even* if $f(-x) = f(x)$ and *odd* if $f(-x) = -f(x)$. Because cos is even and sin is odd, the coefficient integrals over the symmetric interval $[-L, L]$ simplify dramatically.

Proposition 1.4. *If f is even, then $b_n = 0$ for all n, and $a_n = \frac{2}{L}\int_0^L f(x)\cos\frac{n\pi x}{L}\,dx$, so the Fourier series of f contains only cosines. If f is odd, then $a_n = 0$ for all n, and $b_n = \frac{2}{L}\int_0^L f(x)\sin\frac{n\pi x}{L}\,dx$, so the series contains only sines.*

Proof. If f is even, then $f(x)\sin(n\pi x/L)$ is odd (even times odd), and its integral over $[-L, L]$ vanishes; this gives $b_n = 0$. Meanwhile $f(x)\cos(n\pi x/L)$ is even, so its integral over $[-L, L]$ is twice the integral over $[0, L]$. The odd case is symmetric. □

This has a practical consequence. A function defined only on $[0, L]$ can be expanded in *either* a pure sine series or a pure cosine series, by first extending it to $[-L, L]$ as an odd or an even function respectively and then computing the Fourier series of the extension. These are the *half-range expansions.* The sine series builds a function that is zero at $x = 0$ and $x = L$; the cosine series builds one whose derivative is zero there. Which is appropriate depends on the boundary conditions of the problem at hand—a point that becomes important when we solve partial differential equations on an interval.

1.6 Two worked examples

Example 1.5 (The square wave)**.** Let f be the $2L$-periodic odd function with

$$f(x) = \begin{cases} +1, & 0 < x < L, \\ -1, & -L < x < 0. \end{cases}$$

Because f is odd, Proposition 1.4 gives $a_n = 0$ and

$$b_n = \frac{2}{L}\int_0^L \sin\frac{n\pi x}{L}\,dx = \frac{2}{L}\cdot\frac{L}{n\pi}\Big(1-\cos n\pi\Big) = \frac{2}{n\pi}\big(1-(-1)^n\big).$$

Thus $b_n = 0$ for even n and $b_n = 4/(n\pi)$ for odd n, and

$$f(x) = \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{1}{2k+1}\sin\frac{(2k+1)\pi x}{L} = \frac{4}{\pi}\left(\sin\frac{\pi x}{L} + \frac{1}{3}\sin\frac{3\pi x}{L} + \frac{1}{5}\sin\frac{5\pi x}{L} + \cdots\right). \tag{1.6}$$

Only the odd harmonics appear, and their amplitudes decay like $1/n$—slowly, which is the analytic fingerprint of the jump discontinuities, as we will see.

Example 1.6 (The triangular wave)**.** Let g be the $2L$-periodic even extension of $g(x) = |x|$ on $[-L, L]$. By Proposition 1.4 the series is a cosine series with $a_0 = \frac{2}{L}\int_0^L x\,dx = L$ and, integrating by parts,

$$a_n = \frac{2}{L}\int_0^L x\cos\frac{n\pi x}{L}\,dx = \frac{2L}{n^2\pi^2}\big((-1)^n - 1\big),$$

so $a_n = 0$ for even n and $a_n = -4L/(n^2\pi^2)$ for odd n. Hence

$$|x| = \frac{L}{2} - \frac{4L}{\pi^2}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}\cos\frac{(2k+1)\pi x}{L}. \tag{1.7}$$

Now the coefficients decay like $1/n^2$: faster than for the square wave, because g is continuous (only its derivative jumps). The general principle is visible already: *the smoother the function, the faster its Fourier coefficients decay.* Setting $x = 0$ in (1.7) gives the pretty identity $\sum_{k\geq 0}(2k+1)^{-2} = \pi^2/8$.

1.7 Convergence of Fourier series*

Whether the series (1.1) represents f, and in what sense, is settled by the following theorem of Dirichlet, which suffices for every function arising in this course.

Theorem 1.7 (Dirichlet)**.** *Let f be $2L$-periodic and piecewise smooth. Then the Fourier series (1.1) converges at every point x. Where f is continuous, the sum of the series equals $f(x)$. At a point x_0 where f has a jump, the series converges to the average of the one-sided limits,*

$$\frac{f(x_0^+) + f(x_0^-)}{2}.$$

Two features of the theorem account for phenomena visible in any plot of a partial sum.

First, the convergence-to-the-midpoint at a jump is exactly what one should expect from the symmetry of the coefficient integrals: the series cannot "know" which one-sided value to choose, and it splits the difference. For the square wave of Example 1.5, at $x = 0$ every term of (1.6) is zero, so the series sums to 0, which is indeed the midpoint of the jump from -1 to $+1$.

Second, the rate of decay of the coefficients controls the smoothness of the sum and the speed of convergence. We saw $1/n$ decay for the discontinuous square wave and $1/n^2$ for the continuous triangular wave; in general, the more derivatives f has, the faster the coefficients decay and the more rapidly the partial sums converge. This is one of the most useful rules of thumb in the subject and we will return to it for the Fourier transform.

1.8 The Gibbs phenomenon

Theorem 1.7 guarantees *pointwise* convergence, but the manner in which the partial sums approach a function with a jump is surprising. Define the N-th partial sum

$$S_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N} \left(a_n \cos \frac{n\pi x}{L} + b_n \sin \frac{n\pi x}{L} \right).$$

Near a jump, S_N does not merely approximate the step; it *overshoots* it, with a spike just before the corner and a matching dip just after. As N increases the spike narrows and crowds toward the discontinuity, but—and this is the surprise—its height does not shrink to zero. The overshoot settles at approximately 9% of the size of the jump (more precisely a fraction $\frac{1}{\pi}\int_0^\pi \frac{\sin t}{t}\,dx - \frac{1}{2} \approx 0.0895$ of the jump on each side). This persistent overshoot is the *Gibbs phenomenon*.

The resolution of the apparent paradox is the distinction between pointwise and uniform convergence. At any *fixed* point away from the jump, $S_N(x) \to f(x)$, consistent with Theorem 1.7. But the *location* of the worst overshoot moves toward the jump as N grows, so it is never pinned to a single point; the convergence is not uniform in a neighborhood of the discontinuity. The practical lesson is that truncating a Fourier series near a discontinuity introduces ripples that cannot be removed by taking more terms, only relocated—a fact of real consequence in signal processing and numerical analysis.

1.9 Parseval's theorem

Our final result of the chapter expresses the "energy" of a periodic function—the integral of its square—in terms of its Fourier coefficients. It is the periodic ancestor of a relation we will meet again for the Fourier transform and, abstractly, for any orthonormal expansion.

Theorem 1.8 (Parseval)**.** *If f is $2L$-periodic and square-integrable on $[-L, L]$, then*

$$\frac{1}{2L}\int_{-L}^{L} |f(x)|^2\,dx = \sum_{n=-\infty}^{\infty} |c_n|^2 = \left(\frac{a_0}{2}\right)^2 + \frac{1}{2}\sum_{n=1}^{\infty}\left(a_n^2 + b_n^2\right). \tag{1.8}$$

Proof sketch. Write f as its complex series (1.3), form $|f|^2 = f\overline{f}$, and integrate over $[-L, L]$. Expanding the product gives a double sum of integrals $\int_{-L}^{L} e^{im\pi x/L}\overline{e^{in\pi x/L}}\,dx$, and by the orthogonality relation (1.4) only the diagonal terms $m = n$ survive, each contributing $2L\,|c_n|^2$. Dividing by $2L$ gives the first equality; the second follows from the dictionary (1.5). □

Parseval's theorem is useful in two directions. Read from left to right, it computes a sometimes-difficult integral as a sum of squares of coefficients. Read from right to left, it evaluates infinite sums of squares. Applied to the triangular wave of Example 1.6, for instance, it produces the value of $\sum_{k\geq 0}(2k+1)^{-4}$; applied to the square wave it gives $\sum_{k\geq 0}(2k+1)^{-2}=\pi^2/8$. Exercise 1.5 asks you to carry out the second of these.

Exercises

Exercise 1.1. Show that the product of an even function and an odd function is odd, and that the product of two odd functions is even. Use this to give a one-line reason why the Fourier series of an odd function contains no cosine terms.

Exercise 1.2. (Sawtooth.) Compute the Fourier series of the $2L$-periodic odd extension of $f(x)=x$ on $(-L,L)$. Show that $x=\dfrac{2L}{\pi}\displaystyle\sum_{n=1}^{\infty}\frac{(-1)^{n+1}}{n}\sin\frac{n\pi x}{L}$, and explain why the coefficients decay only like $1/n$ even though $f(x)=x$ is smooth on the open interval.

Exercise 1.3. Verify directly from (1.2) that the constant term $a_0/2$ equals the average value $\frac{1}{2L}\int_{-L}^{L}f(x)\,dx$ of f over one period.

Exercise 1.4. (A sum from the boundary value.) Compute the cosine series of the $2L$-periodic even extension of $f(x)=x^2$ on $[-L,L]$. By evaluating the series at a suitable point, deduce the classical identity $\displaystyle\sum_{n=1}^{\infty}\frac{1}{n^2}=\frac{\pi^2}{6}$.

Exercise 1.5. (Parseval.) Apply Theorem 1.8 to the square wave of Example 1.5 and deduce that $\displaystyle\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}=\frac{\pi^2}{8}$.

Exercise 1.6. (Half-range expansions.) For $f(x)=1$ on $[0,L]$, write down both the half-range sine series and the half-range cosine series. Sketch the periodic function that each series represents on $[-3L,3L]$, and state the value each series converges to at $x=0$ and at $x=L$.

Exercise 1.7. *(Convergence at a jump.) Using Theorem 1.7, explain carefully why the square-wave series (1.6) converges to 0 at $x=0$ and at $x=\pm L$, and to ± 1 at interior points. Why does this not contradict the persistence of the Gibbs overshoot?

Chapter 2

The Fourier Transform

2.1 From Fourier series to the Fourier transform

In Chapter 1 we resolved a function of period $2L$ into a discrete superposition of harmonics $\mathrm{e}^{\mathrm{i}n\pi x/L}$, with frequencies $k_n = n\pi/L$ spaced $\Delta k = \pi/L$ apart. A function defined on all of $\mathbb{R}$ that does not repeat has no fundamental period to organize such a sum. We can, however, picture it as a limit of periodic functions whose period $2L$ is sent to infinity: as L grows the allowed frequencies k_n crowd ever closer together, the spacing $\Delta k \to 0$, and the discrete spectrum melts into a continuum. Carrying out that limit replaces the Fourier *series* by the Fourier *transform* and the sum over n by an integral over a continuous frequency variable. The transform is the natural instrument for problems on an unbounded domain, exactly as Fourier series are for problems on a bounded interval.

To watch the limit take shape, write the complex Fourier series of a function f supported in $[-L, L]$,

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \,\mathrm{e}^{\mathrm{i}k_n x}, \qquad c_n = \frac{1}{2L}\int_{-L}^{L} f(y)\,\mathrm{e}^{-\mathrm{i}k_n y}\,dy, \qquad k_n = \frac{n\pi}{L},$$

and substitute the coefficients into the series, writing $1/2L = \Delta k/2\pi$ (recall $\Delta k = \pi/L$):

$$f(x) = \sum_{n=-\infty}^{\infty} \frac{\Delta k}{2\pi}\left(\int_{-L}^{L} f(y)\,\mathrm{e}^{-\mathrm{i}k_n y}\,dy\right)\mathrm{e}^{\mathrm{i}k_n x}.$$

As $L \to \infty$ the bracket becomes a function of the continuous variable k, the spacing $\Delta k \to 0$, and the sum turns into a Riemann integral:

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} f(y)\,\mathrm{e}^{-\mathrm{i}ky}\,dy\right)\mathrm{e}^{\mathrm{i}kx}\,dk. \tag{2.1}$$

The inner integral depends on k alone; it is the object we now name.

Definition 2.1. The *Fourier transform* of an (absolutely integrable) function $f : \mathbb{R} \to \mathbb{C}$ is

$$\widetilde{f}(k) = \mathcal{F}[f](k) := \int_{-\infty}^{\infty} f(x)\,\mathrm{e}^{-\mathrm{i}kx}\,dx, \tag{2.2}$$

and the *inverse Fourier transform* reconstructs f from its transform by

$$f(x) = \mathcal{F}^{-1}[\widetilde{f}](x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \widetilde{f}(k)\,\mathrm{e}^{\mathrm{i}kx}\,dk. \tag{2.3}$$

The pair (2.2)–(2.3) is precisely (2.1) read in the two directions. We call $\widetilde{f}(k)$ the *spectrum* of f; the quantity $\left|\widetilde{f}(k)\right|^2$ measures how much of f is carried by frequencies near k, in a sense made exact by Parseval's theorem in Section 2.9.

Remark 2.2 (Conventions)**.** Several conventions for the constants in (2.2)–(2.3) are in common use: the factor 2π may be placed entirely on the inverse transform, as we do here following McQuarrie; split symmetrically as $1/\sqrt{2\pi}$ on each; or removed by using the frequency variable $\nu = k/2\pi$. Formulas such as Parseval's theorem carry different constants under each choice, so one must fix a convention and keep it. Ours puts the single factor $1/2\pi$ on the inverse transform—exactly the factor that survived the limit above.

2.2 Integral transforms and kernels

The Fourier transform is one member of a large family. An *integral transform* sends a function f to a new function

$$(Tf)(s) = \int_a^b K(s,x)\, f(x)\, dx,$$

where the fixed function $K(s,x)$ is the *kernel*. For the Fourier transform the kernel is $K(k,x) = \mathrm{e}^{-\mathrm{i}kx}$ and the interval is all of $\mathbb{R}$. A second transform we shall need is the *Laplace transform*

$$\mathcal{L}[f](s) = \int_0^\infty \mathrm{e}^{-sx} f(x)\, dx,$$

with kernel e^{-sx} on $[0,\infty)$, which is suited to initial-value problems on the half-line. The reason transforms are useful is always the same: a well-chosen kernel turns the operations appearing in a problem—above all differentiation—into algebra, so that a differential equation for f becomes an algebraic (or lower-order) equation for Tf. We make this precise for the Fourier transform in Section 2.6 and put it to work throughout Chapter 3.

2.3 First examples

Example 2.3 (The rectangular pulse)**.** Let f be the pulse of width $2a$,

$$f(x) = \begin{cases} 1, & |x| < a, \\ 0, & |x| > a. \end{cases}$$

Then

$$\widetilde{f}(k) = \int_{-a}^{a} \mathrm{e}^{-\mathrm{i}kx}\, dx = \frac{\mathrm{e}^{-\mathrm{i}ka} - \mathrm{e}^{\mathrm{i}ka}}{-\mathrm{i}k} = \frac{2\sin ka}{k} = 2a\,\mathrm{sinc}(ka),$$

where $\mathrm{sinc}(u) := \sin u/u$ with $\mathrm{sinc}(0) = 1$. The spectrum is a slowly decaying oscillation whose central lobe occupies $|k| < \pi/a$. Note the *reciprocity of widths*: a narrow pulse (small a) has a broad spectrum, and vice versa. This is the first hint of the uncertainty principle of Section 2.7.

Example 2.4 (The Gaussian)**.** Let $f(x) = \mathrm{e}^{-\alpha x^2}$ with $\alpha > 0$. Completing the square in the exponent,

$$-\alpha x^2 - \mathrm{i}kx = -\alpha\left(x + \frac{\mathrm{i}k}{2\alpha}\right)^2 - \frac{k^2}{4\alpha},$$

so that, shifting the contour of integration (legitimate because the integrand is entire and decays rapidly),

$$\widetilde{f}(k) = \mathrm{e}^{-k^2/4\alpha} \int_{-\infty}^{\infty} \mathrm{e}^{-\alpha(x+\mathrm{i}k/2\alpha)^2}\,dx = \sqrt{\frac{\pi}{\alpha}}\,\mathrm{e}^{-k^2/4\alpha}.$$

The Fourier transform of a Gaussian is a Gaussian. A narrow bell (large α) transforms to a broad one (its width in k is $\sim\sqrt{\alpha}$), again reciprocal widths. Taking $\alpha = \frac{1}{2}$ gives $\mathcal{F}[\mathrm{e}^{-x^2/2}] = \sqrt{2\pi}\,\mathrm{e}^{-k^2/2}$: up to the constant, the standard Gaussian is its own transform. We will meet this function as the heat kernel in Chapter 3 and as the minimizer of the uncertainty product.

Example 2.5 (Decaying exponentials)**.** For the one-sided exponential $f(x) = \mathrm{e}^{-ax}$ for $x > 0$ and 0 for $x < 0$ (with $a > 0$),

$$\widetilde{f}(k) = \int_0^{\infty} \mathrm{e}^{-(a+\mathrm{i}k)x}\,dx = \frac{1}{a+\mathrm{i}k}.$$

For the two-sided exponential $g(x) = \mathrm{e}^{-a|x|}$,

$$\widetilde{g}(k) = \int_{-\infty}^{0} \mathrm{e}^{(a-\mathrm{i}k)x}\,dx + \int_0^{\infty} \mathrm{e}^{-(a+\mathrm{i}k)x}\,dx = \frac{1}{a-\mathrm{i}k} + \frac{1}{a+\mathrm{i}k} = \frac{2a}{a^2+k^2}.$$

The transform is a *Lorentzian.* The pair $\mathrm{e}^{-a|x|} \leftrightarrow 2a/(a^2+k^2)$ will reappear when we solve Laplace's equation in a half-plane.

2.4 The Dirac delta function

Many of the cleanest formulas of the subject involve an object that is not a function in the ordinary sense. The *Dirac delta* δ is a *generalized function*, defined not by its values but by how it behaves inside an integral: for every function φ continuous at a,

$$\int_{-\infty}^{\infty} \delta(x-a)\,\varphi(x)\,dx = \varphi(a). \tag{2.4}$$

This is the *sifting property*: integrating against $\delta(x-a)$ samples the test function φ at the single point $x = a$. The delta is the continuous counterpart of the Kronecker delta of Chapter 1.

Intuitively $\delta(x)$ is an infinitely tall, infinitely narrow spike at the origin enclosing unit area. This picture is made precise by realizing δ as a limit of ordinary bumps of unit area whose width shrinks to zero—for instance the narrowing Gaussians $\delta_\varepsilon(x) = (4\pi\varepsilon)^{-1/2}\mathrm{e}^{-x^2/4\varepsilon}$, or the boxes of height $1/2\varepsilon$ on $(-\varepsilon, \varepsilon)$. For any such family, $\int \delta_\varepsilon\,\varphi \to \varphi(0)$ as $\varepsilon \to 0$, which is (2.4).

The defining property (2.4) dictates the algebra of the delta. The substitution rules give that δ is even, $\delta(-x) = \delta(x)$, and that it rescales as

$$\delta(ax) = \frac{1}{|a|}\,\delta(x), \qquad x\,\delta(x) = 0.$$

Its derivative δ' is defined, consistently with integration by parts, by $\int \delta'(x-a)\varphi(x)\,dx = -\varphi'(a)$, from which $x\delta'(x) = -\delta(x)$.

Remark 2.6. ⋆The rigorous home of the delta is the theory of *distributions*, in which δ is a continuous linear functional on a space of smooth, rapidly decaying test functions, and (2.4) is its definition. Every formal manipulation below can be justified by replacing δ with an approximating family δ_ε, computing, and letting $\varepsilon \to 0$ at the end.

2.5 The transform of the delta and the integral representation of δ

Applying the definition (2.2) and the sifting property,

$$\mathcal{F}[\delta](k) = \int_{-\infty}^{\infty} \delta(x)\, \mathrm{e}^{-\mathrm{i}kx}\, dx = 1.$$

The transform of a unit spike is the constant function 1: a perfectly localized signal contains every frequency in equal measure. Inverting via (2.3) gives the single most useful identity of Fourier analysis,

$$\delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}kx}\, dk. \tag{2.5}$$

This *integral representation of the delta* is the continuum limit of the orthogonality of the exponentials in Chapter 1, and it is the engine that drives Fourier inversion. Read in the other direction, it says $\mathcal{F}[1](k) = 2\pi\,\delta(k)$. More generally,

$$\mathcal{F}[\mathrm{e}^{\mathrm{i}k_0 x}](k) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{i}(k_0-k)x}\, dx = 2\pi\,\delta(k-k_0),$$

so a pure oscillation has a spectrum concentrated at a single frequency, as it must. Splitting cosine and sine into exponentials,

$$\mathcal{F}[\cos k_0 x] = \pi\big[\delta(k-k_0) + \delta(k+k_0)\big], \qquad \mathcal{F}[\sin k_0 x] = -\mathrm{i}\pi\big[\delta(k-k_0) - \delta(k+k_0)\big].$$

2.6 Operational properties

The usefulness of the transform rests on a short list of rules, each proved by a one-line manipulation of the integral (2.2).

Proposition 2.7. *Let f, g have Fourier transforms $\widetilde{f}, \widetilde{g}$, and let $a \in \mathbb{R}$, $a \neq 0$.*

(i) (Linearity) $\mathcal{F}[\alpha f + \beta g] = \alpha \widetilde{f} + \beta \widetilde{g}$.

(ii) (Translation) $\mathcal{F}[f(x-a)](k) = \mathrm{e}^{-\mathrm{i}ka}\, \widetilde{f}(k)$.

(iii) (Modulation) $\mathcal{F}[\mathrm{e}^{\mathrm{i}k_0 x} f(x)](k) = \widetilde{f}(k-k_0)$.

(iv) (Scaling) $\mathcal{F}[f(ax)](k) = \dfrac{1}{|a|}\, \widetilde{f}\left(\dfrac{k}{a}\right)$.

(v) (Derivative) *If f and its derivatives vanish at $\pm\infty$, then* $\mathcal{F}[f^{(n)}](k) = (\mathrm{i}k)^n\, \widetilde{f}(k)$.

(vi) (Multiplication by x) $\mathcal{F}[x^n f(x)](k) = \mathrm{i}^n\, \dfrac{d^n \widetilde{f}}{dk^n}(k)$.

Proof. (i) is the linearity of the integral. For (ii), substitute $u = x - a$ in $\int f(x-a)\mathrm{e}^{-\mathrm{i}kx}\, dx$; the change produces the factor $\mathrm{e}^{-\mathrm{i}ka}$. Property (iii) is the same computation with the roles of x and k exchanged. For (iv), substitute $u = ax$ and track the Jacobian $1/|a|$ and the sign of a. For the derivative rule (v), integrate by parts once,

$$\int_{-\infty}^{\infty} f'(x)\mathrm{e}^{-\mathrm{i}kx}\, dx = \big[f(x)\mathrm{e}^{-\mathrm{i}kx}\big]_{-\infty}^{\infty} + \mathrm{i}k \int_{-\infty}^{\infty} f(x)\mathrm{e}^{-\mathrm{i}kx}\, dx = \mathrm{i}k\, \widetilde{f}(k),$$

the boundary term vanishing by hypothesis; iterate for $f^{(n)}$. Finally (vi) follows by differentiating (2.2) under the integral sign, since $\partial_k \mathrm{e}^{-\mathrm{i}kx} = -\mathrm{i}x\, \mathrm{e}^{-\mathrm{i}kx}$. □

Property (v) is the one that matters most: *differentiation in x becomes multiplication by* $\mathrm{i}k$. It is what allows the transform to convert a linear constant-coefficient differential equation into an algebraic one, the theme of the next chapter.

2.7 The wave train and the uncertainty principle

Consider a finite wave train, a pure oscillation switched on for a finite length,

$$f(x) = \begin{cases} \mathrm{e}^{\mathrm{i}k_0 x}, & |x| < L, \\ 0, & |x| > L. \end{cases}$$

Its transform combines the modulation and pulse computations above:

$$\widetilde{f}(k) = \int_{-L}^{L} \mathrm{e}^{\mathrm{i}(k_0 - k)x}\, dx = \frac{2\sin\big((k_0 - k)L\big)}{k_0 - k}.$$

The spectrum is peaked at $k = k_0$ with a central lobe of half-width π/L. As the train lengthens, $L \to \infty$, the lobe sharpens and $\widetilde{f} \to 2\pi\,\delta(k - k_0)$: an infinitely long oscillation is perfectly monochromatic. A short train, by contrast, is spread over a wide band. The product of the spatial extent ($\sim 2L$) and the spectral width ($\sim \pi/L$) is independent of L: *one cannot make a signal both brief and pure in frequency.* This reciprocity has a sharp quantitative form.

Theorem 2.8 (Uncertainty principle)**.** **Let f be square-integrable with $\int x\,|f|^2\,dx = 0$ and $\int k\,\left|\widetilde{f}\right|^2\,dk = 0$ (zero spatial and spectral means; otherwise re-center), and define the spreads*

$$(\Delta x)^2 = \frac{\int x^2\,|f(x)|^2\,dx}{\int |f(x)|^2\,dx}, \qquad (\Delta k)^2 = \frac{\int k^2\left|\widetilde{f}(k)\right|^2\,dk}{\int \left|\widetilde{f}(k)\right|^2\,dk}.$$

Then $\Delta x\,\Delta k \geq \frac{1}{2}$, with equality if and only if f is a Gaussian.

Proof. By Plancherel's theorem (Section 2.9) and the derivative rule, $\int k^2\left|\widetilde{f}\right|^2\,dk = 2\pi\int |f'|^2\,dx$ and $\int \left|\widetilde{f}\right|^2\,dk = 2\pi\int |f|^2\,dx$, so $(\Delta k)^2 = \int |f'|^2\,dx / \int |f|^2\,dx$. Differentiating $x\,|f|^2$ and integrating over $\mathbb{R}$ (the boundary terms vanish) gives $\int |f|^2\,dx = -2\int x\,\mathrm{Re}\big(\overline{f}\,f'\big)\,dx$, whence

$$\tfrac{1}{2}\int |f|^2\,dx = \left|\mathrm{Re}\int x\,\overline{f}\,f'\,dx\right| \leq \left|\int x\,\overline{f}\,f'\,dx\right| \leq \Big(\int x^2\,|f|^2\,dx\Big)^{1/2}\Big(\int |f'|^2\,dx\Big)^{1/2},$$

the last step by the Cauchy–Schwarz inequality. Squaring and dividing by $\big(\int |f|^2\big)^2$ yields $(\Delta x)^2(\Delta k)^2 \geq \frac{1}{4}$. Equality in Cauchy–Schwarz forces $f' = -c\,x f$ for a constant $c > 0$, i.e. a Gaussian. □

In quantum mechanics a particle's state is described by a wavefunction $\psi(x)$ whose squared modulus is the probability density for position, while $\left|\widetilde{\psi}(k)\right|^2$ gives the density for momentum $p = \hbar k$. Theorem 2.8 becomes $\Delta x\,\Delta p \geq \hbar/2$: Heisenberg's uncertainty relation is a theorem about Fourier transforms.

2.8 Convolution

Definition 2.9. The *convolution* of f and g is

$$(f * g)(x) = \int_{-\infty}^{\infty} f(y)\, g(x-y)\, dy.$$

Convolution is the operation of *blurring* f by the window g: each value of $f * g$ is a weighted average of nearby values of f, the weights given by a flipped, shifted copy of g. Equivalently, $f * g$ is the response of a linear, translation-invariant system—an optical instrument, an electrical filter—to the input f, where g is the system's response to a unit impulse δ. The operation is commutative and associative (substitute $u = x - y$ to see $f * g = g * f$). Its interaction with the transform is the cornerstone of the next chapter.

Theorem 2.10 (Convolution theorem)**.** *$\mathcal{F}[f * g] = \widetilde{f}\,\widetilde{g}$. Dually, $\mathcal{F}[fg] = \frac{1}{2\pi}\,\widetilde{f} * \widetilde{g}$.*

Proof. Interchanging the order of integration,

$$\mathcal{F}[f * g](k) = \iint f(y) g(x-y)\, \mathrm{e}^{-\mathrm{i}kx}\, dy\, dx = \int f(y) \mathrm{e}^{-\mathrm{i}ky} \left(\int g(x-y) \mathrm{e}^{-\mathrm{i}k(x-y)}\, dx \right) dy = \widetilde{f}(k)\, \widetilde{g}(k).$$

The dual statement follows by writing f and g through the inversion formula (2.3) and using (2.5). □

Example 2.11. Convolving the pulse of Example 2.3 with itself gives, geometrically, the length of the overlap of two boxes offset by x:

$$(f * f)(x) = \begin{cases} 2a - |x|\,, & |x| < 2a, \\ 0, & |x| > 2a, \end{cases}$$

a triangle of base $4a$. Consistently with Theorem 2.10, its transform is $\widetilde{f}(k)^2 = (2 \sin ka/k)^2$. More generally, convolving two Gaussians yields another Gaussian: $\mathrm{e}^{-\alpha x^2} * \mathrm{e}^{-\beta x^2} = \sqrt{\pi/(\alpha+\beta)}\, \mathrm{e}^{-\alpha\beta x^2/(\alpha+\beta)}$ (Exercise).

2.9 Parseval–Plancherel

Theorem 2.12 (Parseval–Plancherel)**.** *For square-integrable f, g,*

$$\int_{-\infty}^{\infty} f(x)\, \overline{g(x)}\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{f}(k)\, \overline{\widetilde{g}(k)}\, dk,$$

and in particular $\displaystyle\int_{-\infty}^{\infty} |f(x)|^2\, dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} \left| \widetilde{f}(k) \right|^2 dk.$

Proof sketch. Write $\overline{g(x)}$ through the conjugate of the inversion formula (2.3), insert into the left-hand side, and exchange the order of integration; the inner x-integral produces $\widetilde{f}(k)$, leaving $\frac{1}{2\pi} \int \widetilde{f}\, \overline{\widetilde{g}}\, dk$. □

The identity says that, up to the factor $1/2\pi$, the transform preserves the "energy" $\int |f|^2$. Read from right to left it evaluates integrals: for the pulse of Example 2.3 with $a = 1$, $\int |f|^2\, dx = 2$ while $\widetilde{f}(k) = 2 \sin k/k$, so Theorem 2.12 gives the elegant value

$$\int_{-\infty}^{\infty} \left(\frac{\sin k}{k} \right)^2 dk = \pi.$$

In quantum mechanics the same identity is the statement that total probability is the same whether computed in position or in momentum space.

Exercises

Exercise 2.1. Compute the Fourier transform of $f(x) = \mathrm{e}^{-a|x|}$ directly, and confirm it equals $2a/(a^2 + k^2)$. Use the inversion formula to write $\mathrm{e}^{-a|x|}$ as an integral, and check the value at $x = 0$.

Exercise 2.2. Prove the scaling rule, Proposition 2.7(iv), and use it together with the transform of a unit box to obtain the transform of the box of width $2a$ without integrating again.

Exercise 2.3. Carry out the completing-the-square computation of Example 2.4 in full detail, and verify the width product $(\Delta x)(\Delta k) = \frac{1}{2}$ for the Gaussian, confirming that it saturates Theorem 2.8.

Exercise 2.4. Establish the delta identities $\delta(ax) = \delta(x)/\,|a|$ and $x\,\delta'(x) = -\delta(x)$ from the defining property (2.4) and the definition of δ'.

Exercise 2.5. Verify by direct integration that the convolution of the box with itself is the triangle of Example 2.11, and interpret the result as an overlap area.

Exercise 2.6. *Use the Fourier transform to solve $-u''(x) + a^2u(x) = f(x)$ on $\mathbb{R}$ (with $a > 0$ and $u \to 0$ at $\pm\infty$). Show that $\widetilde{u}(k) = \widetilde{f}(k)/(k^2 + a^2)$ and hence $u = G * f$ with Green's function $G(x) = \mathrm{e}^{-a|x|}/(2a)$. Write the solution as an explicit integral.

Exercise 2.7. Apply Parseval's theorem to the box of width 2 to evaluate $\int_{-\infty}^{\infty} (\sin k/k)^2\, dk$, and confirm the value π.

Chapter 3

Fourier Transforms and Partial Differential Equations

3.1 The strategy

Three of the central equations of mathematical physics—the heat, wave, and Laplace equations—are linear with constant coefficients. On the whole line the Fourier transform reduces each of them to a problem solvable almost by inspection. The mechanism is the derivative rule, Proposition 2.7(v): because $\mathcal{F}[\partial_x f] = \mathrm{i}k\,\widetilde{f}$, a spatial derivative becomes multiplication by $\mathrm{i}k$, so a partial differential equation in x and t collapses into an *ordinary* differential equation in t for the transform $\widetilde{u}(k,t)$, with the frequency k entering only as a parameter. We solve that ODE and then invert. The inversion almost always takes the form of a *convolution* of the data with a fixed function, the *fundamental solution*, by the convolution theorem. The same three steps—transform, solve the ODE in t, invert—recur in every example below.

3.2 The heat equation on the line

The temperature $u(x,t)$ in an infinite bar with initial profile f obeys

$$u_t = \kappa\, u_{xx} \quad (x \in \mathbb{R},\ t > 0), \qquad u(x,0) = f(x),$$

where $\kappa > 0$ is the diffusivity. Transform in x, writing $\widetilde{u}(k,t) = \int u(x,t)\mathrm{e}^{-\mathrm{i}kx}\,dx$; the rule $\mathcal{F}[u_{xx}] = -k^2\widetilde{u}$ turns the equation into the ordinary differential equation

$$\widetilde{u}_t(k,t) = -\kappa k^2\,\widetilde{u}(k,t), \qquad \widetilde{u}(k,0) = \widetilde{f}(k),$$

whose solution is $\widetilde{u}(k,t) = \widetilde{f}(k)\,\mathrm{e}^{-\kappa k^2 t}$. From the Gaussian transform of Example 2.4 (read backwards, with $\alpha = 1/4\kappa t$), $\mathrm{e}^{-\kappa k^2 t}$ is the transform of the *heat kernel*

$$\Phi(x,t) = \frac{1}{\sqrt{4\pi\kappa t}}\,\mathrm{e}^{-x^2/4\kappa t}. \tag{3.1}$$

By the convolution theorem,

$$u(x,t) = \big(\Phi(\cdot,t) * f\big)(x) = \frac{1}{\sqrt{4\pi\kappa t}}\int_{-\infty}^{\infty} \mathrm{e}^{-(x-y)^2/4\kappa t}\,f(y)\,dy. \tag{3.2}$$

For the point initial condition $f = \delta(x-a)$ the solution is simply $u(x,t) = \Phi(x-a,t)$, the spreading bell of heat released from a point.

The formula (3.2) encodes the physics of diffusion. First, *instantaneous smoothing*: for any $t > 0$ the kernel is infinitely differentiable, so $u(\cdot,t)$ is smooth even when the initial data f has corners or jumps. Second, *infinite propagation speed*: because Φ is positive everywhere, heat initially confined near a point is felt—if only exponentially weakly—at every distance the instant $t > 0$. This is an idealization; no signal truly travels infinitely fast, and it stands in sharp contrast to the wave equation below. Third, *conservation of heat*: $\int u\,dx = \int f\,dx$ for all t, since $\int \Phi\,dx = 1$. Fourth, *spreading and decay*: the profile widens like $\sqrt{t}$ and its height falls like $1/\sqrt{t}$, the signature of a random walk.

3.3 The wave equation on the line

A disturbance $u(x,t)$ on an infinite string obeys

$$u_{tt} = c^2 u_{xx}, \qquad u(x,0) = f(x), \quad u_t(x,0) = g(x),$$

with wave speed c. Transforming in x gives the harmonic-oscillator equation $\widetilde{u}_{tt} = -c^2k^2\widetilde{u}$ for each k, with general solution $\widetilde{u}(k,t) = A(k)\cos(ckt) + B(k)\sin(ckt)$. The initial conditions fix $A = \widetilde{f}$ and $ckB = \widetilde{g}$, so

$$\widetilde{u}(k,t) = \widetilde{f}(k)\cos(ckt) + \widetilde{g}(k)\,\frac{\sin(ckt)}{ck}.$$

To invert the first term, write $\cos(ckt) = \frac{1}{2}(\mathrm{e}^{\mathrm{i}ckt} + \mathrm{e}^{-\mathrm{i}ckt})$; by the translation rule, $\mathcal{F}^{-1}[\mathrm{e}^{\pm \mathrm{i}ckt}\widetilde{f}] = f(x \pm ct)$, so this term inverts to $\frac{1}{2}[f(x+ct) + f(x-ct)]$. The second term involves a division by k, which corresponds to integration; one finds that it inverts to $(1/2c)\int_{x-ct}^{x+ct} g(s)\,ds$. Together these give *d'Alembert's formula*

$$u(x,t) = \tfrac{1}{2}\big[f(x-ct) + f(x+ct)\big] + \frac{1}{2c}\int_{x-ct}^{x+ct} g(s)\,ds. \tag{3.3}$$

The structure of (3.3) is the physics of waves. The displacement is a superposition of two copies of the initial shape, one travelling right and one left at speed c, plus a contribution from the initial velocity. There is *no smoothing*: a corner in f propagates forever as a corner, reflecting the time-reversibility of the wave equation (again unlike heat). And the speed is *finite*: $u(x,t)$ depends only on the data in the interval $[x-ct, x+ct]$, its *domain of dependence*, and the data at a point influence the solution only inside the forward cone, its *domain of influence*. Signals travel at exactly speed c, no faster and no slower.

3.4 Laplace's equation in the upper half-plane

We seek a function $u(x,y)$ harmonic in the upper half-plane and matching given boundary data on the edge,

$$u_{xx} + u_{yy} = 0 \quad (y > 0), \qquad u(x,0) = f(x), \qquad u \text{ bounded as } y \to \infty.$$

Transforming in x turns the equation into $\widetilde{u}_{yy} - k^2\widetilde{u} = 0$, an ODE in y with general solution $A(k)\mathrm{e}^{|k|y} + B(k)\mathrm{e}^{-|k|y}$. Boundedness as $y \to \infty$ forces $A \equiv 0$, and the boundary condition gives $B = \widetilde{f}$, so

$$\widetilde{u}(k,y) = \widetilde{f}(k)\,\mathrm{e}^{-|k|y}.$$

From the Lorentzian pair of Example 2.5, read by duality, $e^{-|k|y}$ is the transform of the *Poisson kernel*

$$P_y(x) = \frac{1}{\pi}\,\frac{y}{x^2+y^2}$$

(this is verified in Exercise 3.4). The convolution theorem then gives *Poisson's integral formula* for the half-plane,

$$u(x,y) = (P_y * f)(x) = \frac{y}{\pi}\int_{-\infty}^{\infty} \frac{f(x')}{(x-x')^2+y^2}\,dx'. \tag{3.4}$$

The harmonic extension of the boundary data is thus a weighted average of those data, the weight being the bell-shaped P_y, which concentrates ever more tightly around $x' = x$ as $y \to 0^+$ (indeed $P_y \to \delta$). This is one face of the *mean-value property* of harmonic functions, and it anticipates the conformal-mapping methods of Part II, where other domains are mapped to the half-plane precisely so that (3.4) can be used.

3.5 The multidimensional Fourier transform

In $\mathbb{R}^n$ the transform and its inverse read

$$\widetilde{f}(\mathbf{k}) = \int_{\mathbb{R}^n} f(\mathbf{x})\,e^{-i\mathbf{k}\cdot\mathbf{x}}\,d^n x, \qquad f(\mathbf{x}) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n} \widetilde{f}(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{x}}\,d^n k.$$

Because the kernel factorizes, $e^{-i\mathbf{k}\cdot\mathbf{x}} = \prod_j e^{-ik_j x_j}$, the transform of a product $f(\mathbf{x}) = \prod_j f_j(x_j)$ is the product of one-dimensional transforms. This makes the heat equation in $\mathbb{R}^n$, $u_t = \kappa\,\Delta u$ with $u(\mathbf{x},0) = f(\mathbf{x})$, no harder than the one-dimensional case: transforming gives $\widetilde{u} = \widetilde{f}\,e^{-\kappa|\mathbf{k}|^2 t}$, and since $e^{-\kappa|\mathbf{k}|^2 t} = \prod_j e^{-\kappa k_j^2 t}$ the fundamental solution is a product of one-dimensional heat kernels,

$$\Phi_n(\mathbf{x},t) = \frac{1}{(4\pi\kappa t)^{n/2}}\,e^{-|\mathbf{x}|^2/4\kappa t}.$$

When the initial data are rotationally symmetric—f a function of $r = |\mathbf{x}|$ only—so is the solution, and the calculation reduces, in two dimensions, to convolution against $\Phi_2(\mathbf{x},t) = (4\pi\kappa t)^{-1}e^{-r^2/4\kappa t}$, literally the product of two one-dimensional kernels.

3.6 Series or transform?

Part I has presented two faces of one idea. On a *bounded* interval with periodic or homogeneous boundary conditions, the natural frequencies form a *discrete* set and a function is built from a Fourier *series* (Chapter 1); on an *unbounded* domain the frequencies form a *continuum* and the building block is the Fourier *transform* (this chapter). The bridge between them is the integral representation (2.5): the relation $\frac{1}{2\pi}\int e^{ikx}\,dk = \delta(x)$ is exactly the continuum limit of the discrete orthogonality of the exponentials. Underlying both is a single principle that will return, in abstract dress, in Part III: *to solve a linear differential equation, expand in the eigenfunctions of its differential operator*—here the exponentials e^{ikx}, which the operator ∂_x^2 multiplies by $-k^2$. Problems on the half-line, intermediate between the two cases, are handled by sine and cosine transforms or by the method of images (Exercise 3.5).

Exercises

Exercise 3.1. Solve the heat equation on $\mathbb{R}$ with initial data $f(x) = \mathrm{e}^{-x^2}$. Show that the solution remains Gaussian for all t, and identify how its width and height change with t.

Exercise 3.2. Fill in the inversion of the second term in the wave calculation to derive d'Alembert's formula (3.3) in full, including the initial-velocity integral. (Hint: division by $\mathrm{i}k$ in the transform domain corresponds to integration in x.)

Exercise 3.3. Verify directly that the heat kernel (3.1) satisfies $\Phi_t = \kappa\Phi_{xx}$ for $t > 0$, and that $\int_{-\infty}^{\infty} \Phi(x,t)\,dx = 1$ for every $t > 0$.

Exercise 3.4. Confirm that the Fourier transform of the Poisson kernel $P_y(x) = (1/\pi)\,y/(x^2 + y^2)$ is $\mathrm{e}^{-|k|y}$, using the exponential pair of Example 2.5 and the symmetry between $\mathcal{F}$ and $\mathcal{F}^{-1}$.

Exercise 3.5. *(Half-line by images.) Solve the heat equation on $x > 0$ with the boundary condition $u(0,t) = 0$ by extending the initial data to an *odd* function on all of $\mathbb{R}$ and applying (3.2). Explain why the odd extension enforces the boundary condition for all t.

Exercise 3.6. Compute the n-dimensional heat kernel by writing the solution of $u_t = \kappa\Delta u$ as a product of one-dimensional solutions, and verify that $\int_{\mathbb{R}^n} \Phi_n(\mathbf{x},t)\,d^n x = 1$.

Chapter 4

Complex Numbers and Elementary Functions

4.1 The complex number system

The equation $x^2 + 1 = 0$ has no real solution, because the square of a real number is never negative. Rather than abandon it, we enlarge the number system by adjoining a single new symbol i with the defining property

$$\mathrm{i}^2 = -1,$$

and decree that it combines with the reals by the ordinary rules of algebra. The result is the system of *complex numbers*, within which not only $x^2 + 1 = 0$ but *every* polynomial equation acquires a solution—a fact (the fundamental theorem of algebra) we shall prove in Section 6.7.

Definition 4.1. A *complex number* is an expression $z = x + \mathrm{i}y$ with $x, y \in \mathbb{R}$. Its *real part* is $\operatorname{Re} z = x$ and its *imaginary part* is $\operatorname{Im} z = y$ (a real number, despite the name). The set of all complex numbers is denoted $\mathbb{C}$, and two complex numbers are equal exactly when their real and imaginary parts agree.

Addition and multiplication are forced by ordinary algebra together with $\mathrm{i}^2 = -1$:

$$(a + \mathrm{i}b) + (c + \mathrm{i}d) = (a + c) + \mathrm{i}(b + d), \qquad (a + \mathrm{i}b)(c + \mathrm{i}d) = (ac - bd) + \mathrm{i}(ad + bc).$$

The *complex conjugate* of $z = x + \mathrm{i}y$ is $\bar{z} = x - \mathrm{i}y$; conjugation reflects z across the real axis and reverses the sign of the imaginary part. A direct computation gives

$$z\bar{z} = (x + \mathrm{i}y)(x - \mathrm{i}y) = x^2 + y^2 \ \geq 0,$$

a nonnegative real number whose square root is the *modulus* $|z| = \sqrt{x^2 + y^2}$. Conjugation turns division into multiplication: for $z \neq 0$,

$$\frac{w}{z} = \frac{w\bar{z}}{z\bar{z}} = \frac{w\bar{z}}{|z|^2}.$$

With these operations $\mathbb{C}$ is a *field* containing $\mathbb{R}$ as the numbers with zero imaginary part.

Remark 4.2. Unlike $\mathbb{R}$, the field $\mathbb{C}$ cannot be ordered in a way compatible with its arithmetic: if it could, then $\mathrm{i}^2 = -1$ would have to be both nonnegative (as a square) and negative (as -1). The symbols "<" and ">" are therefore meaningless between non-real complex numbers; only $|z|$, a real number, may be compared.

4.2 The complex plane, modulus and argument

Writing $z = x + \mathrm{i}y$ as the point (x, y) identifies $\mathbb{C}$ with the Euclidean plane, now called the *complex plane*. The horizontal axis carries the real numbers and the vertical axis the purely imaginary ones. In this picture $|z|$ is the distance from z to the origin, addition is the parallelogram law of vectors, and conjugation is reflection in the real axis. The modulus obeys the familiar inequalities

$$|zw| = |z|\,|w|\,, \qquad |z + w| \le |z| + |w| \text{ (triangle inequality).}$$

A nonzero z is equally well described by its distance $r = |z|$ from the origin and the angle θ its radius vector makes with the positive real axis. This angle is the *argument* of z, written $\arg z$; it satisfies $\tan\theta = y/x$ together with the correct quadrant of (x, y). The argument is determined only up to adding integer multiples of 2π, so $\arg z$ is a *multivalued* quantity. The single value lying in $(-\pi, \pi]$ is the *principal argument*, written $\operatorname{Arg} z$. In terms of r and θ the *polar form* of z is

$$z = r(\cos\theta + \mathrm{i}\sin\theta).$$

4.3 Euler's formula and the polar form

The polar form becomes vastly more useful once the trigonometric factor is recognized as an exponential. Comparing the Taylor series

$$\mathrm{e}^{\mathrm{i}\theta} = \sum_{n=0}^{\infty} \frac{(\mathrm{i}\theta)^n}{n!} = \Big(1 - \tfrac{\theta^2}{2!} + \cdots\Big) + \mathrm{i}\Big(\theta - \tfrac{\theta^3}{3!} + \cdots\Big)$$

with the series for cosine and sine gives *Euler's formula*

$$\mathrm{e}^{\mathrm{i}\theta} = \cos\theta + \mathrm{i}\sin\theta, \tag{4.1}$$

so that every nonzero complex number can be written compactly as $z = r\,\mathrm{e}^{\mathrm{i}\theta}$. (Setting $\theta = \pi$ gives the celebrated identity $\mathrm{e}^{\mathrm{i}\pi} + 1 = 0$.) In this form multiplication is transparent: if $z_1 = r_1\mathrm{e}^{\mathrm{i}\theta_1}$ and $z_2 = r_2\mathrm{e}^{\mathrm{i}\theta_2}$ then

$$z_1 z_2 = r_1 r_2\, \mathrm{e}^{\mathrm{i}(\theta_1+\theta_2)},$$

so *moduli multiply and arguments add.* Geometrically, multiplying by $r\mathrm{e}^{\mathrm{i}\theta}$ is a rotation by θ followed by a scaling by r. From the case $r = 1$ we obtain *de Moivre's theorem*,

$$(\cos\theta + \mathrm{i}\sin\theta)^n = \cos n\theta + \mathrm{i}\sin n\theta,$$

a compact engine for trigonometric identities.

Example 4.3. Expanding $(\cos\theta + \mathrm{i}\sin\theta)^3$ by the binomial theorem and matching real and imaginary parts with $\cos 3\theta + \mathrm{i}\sin 3\theta$ gives at once

$$\cos 3\theta = 4\cos^3\theta - 3\cos\theta, \qquad \sin 3\theta = 3\sin\theta - 4\sin^3\theta.$$

4.4 Roots and powers

Integer powers are immediate in polar form, $z^n = r^n e^{in\theta}$. Roots are more interesting because of the multivaluedness of the argument. To solve $w^n = z$ for $z = re^{i\theta} \neq 0$, write $w = \rho e^{i\phi}$ and match moduli and arguments: $\rho = r^{1/n}$ and $n\phi = \theta + 2\pi k$. As k runs over $0, 1, \ldots, n-1$ this produces n *distinct* roots,

$$w_k = r^{1/n} \exp\left(i\,\frac{\theta + 2\pi k}{n}\right), \qquad k = 0, 1, \ldots, n-1, \tag{4.2}$$

equally spaced by angle $2\pi/n$ around the circle of radius $r^{1/n}$; they are the vertices of a regular n-gon. The roots of $w^n = 1$, the *n-th roots of unity*, are $e^{2\pi i k/n}$.

Example 4.4. The cube roots of unity are 1, $e^{2\pi i/3} = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$, and $e^{4\pi i/3} = -\frac{1}{2} - \frac{\sqrt{3}}{2}i$, forming an equilateral triangle inscribed in the unit circle; their sum is zero, as the sum of all n-th roots of unity always is.

4.5 Multivalued functions and branch cuts

Because $\arg z$ is determined only modulo 2π, any function built from it inherits the ambiguity. We have already seen this in the n values of an n-th root; it will appear again, more seriously, for the logarithm and general powers. The standard remedy is to make a single-valued, continuous choice by restricting the argument to one interval of length 2π and removing a curve—a *branch cut*—across which the chosen value would otherwise jump. The most common choice uses the principal argument $\operatorname{Arg} z$ and cuts the plane along the negative real axis. We develop this for the logarithm in Section 4.7; the points where no continuous single-valued choice is possible in any neighborhood, such as the origin for the logarithm, are *branch points*.

4.6 The exponential and trigonometric functions

Euler's formula (4.1) suggests the definition of the exponential of an arbitrary complex number:

$$e^z = e^{x+iy} := e^x(\cos y + i \sin y).$$

This e^z reduces to the real exponential when $y = 0$, obeys the addition law $e^{z+w} = e^z e^w$, has modulus $|e^z| = e^x$, and is therefore *never zero*. It is, however, *periodic* with the purely imaginary period $2\pi i$, since $e^{2\pi i} = 1$—a feature with no real-variable analogue.

Inverting Euler's formula expresses cosine and sine through exponentials, and the same expressions *define* the trigonometric functions of a complex variable:

$$\cos z = \frac{e^{iz} + e^{-iz}}{2}, \qquad \sin z = \frac{e^{iz} - e^{-iz}}{2i}.$$

They satisfy the usual identities, including $\sin^2 z + \cos^2 z = 1$, and reduce to the real functions on the real axis. The closely related *hyperbolic functions* are $\cosh z = (e^z + e^{-z})/2$ and $\sinh z = (e^z - e^{-z})/2$, and the two families are linked by $\cos(iz) = \cosh z$ and $\sin(iz) = i \sinh z$. One striking departure from the real case: $\cos z$ and $\sin z$ are *unbounded* on $\mathbb{C}$, growing like $e^{|y|}$ as one moves off the real axis (Exercise 4.5).

4.7 The complex logarithm

The logarithm is the inverse of the exponential, so $w = \log z$ should mean $\mathrm{e}^w = z$. Writing $w = u + \mathrm{i}v$ and $z = r\mathrm{e}^{\mathrm{i}\theta}$, the equation $\mathrm{e}^u \mathrm{e}^{\mathrm{i}v} = r\mathrm{e}^{\mathrm{i}\theta}$ forces $\mathrm{e}^u = r$ and $v = \theta + 2\pi k$, that is

$$\log z = \ln|z| + \mathrm{i}\,\arg z = \ln r + \mathrm{i}(\theta + 2\pi k), \qquad k \in \mathbb{Z}. \tag{4.3}$$

The logarithm is therefore *infinitely multivalued*, its values differing by integer multiples of $2\pi\mathrm{i}$—a direct consequence of the periodicity of e^z. Selecting the principal argument gives the *principal logarithm*

$$\operatorname{Log} z = \ln|z| + \mathrm{i}\,\operatorname{Arg} z,$$

a single-valued function on the plane cut along the negative real axis, with a branch point at 0. One must handle it with care: because arguments add only modulo 2π, the identity $\operatorname{Log}(z_1 z_2) = \operatorname{Log} z_1 + \operatorname{Log} z_2$ can fail by $2\pi\mathrm{i}$.

Example 4.5. Since $|-1| = 1$ and $\operatorname{Arg}(-1) = \pi$, we get $\operatorname{Log}(-1) = \mathrm{i}\pi$; similarly $\operatorname{Log}\mathrm{i} = \mathrm{i}\pi/2$. There is nothing paradoxical here once one accepts that the logarithm of a negative number is complex.

4.8 Complex powers

With a logarithm in hand, arbitrary powers are defined by

$$z^a = \mathrm{e}^{a \log z},$$

which is in general multivalued through $\log z$; its *principal value* uses $\operatorname{Log} z$. This definition reproduces the familiar cases (integer a gives the single ordinary power; $a = 1/n$ recovers the n roots of (4.2)) but also yields some surprises.

Example 4.6. Take $z = \mathrm{i}$ and $a = \mathrm{i}$. Then $\log \mathrm{i} = \mathrm{i}(\pi/2 + 2\pi k)$, so

$$\mathrm{i}^{\mathrm{i}} = \mathrm{e}^{\mathrm{i}\log\mathrm{i}} = \mathrm{e}^{\mathrm{i}\cdot\mathrm{i}(\pi/2+2\pi k)} = \mathrm{e}^{-(\pi/2+2\pi k)},$$

a set of *real* numbers, with principal value $\mathrm{e}^{-\pi/2} \approx 0.2079$. A purely imaginary number raised to a purely imaginary power is real.

Exercises

Exercise 4.1. Write $z = 1 + \mathrm{i}$ and $z = -\sqrt{3} + \mathrm{i}$ in polar form $r\mathrm{e}^{\mathrm{i}\theta}$ with $\theta = \operatorname{Arg} z$, and compute z^6 in each case.

Exercise 4.2. Find all four solutions of $z^4 = -16$ and plot them in the complex plane, confirming that they are the vertices of a square.

Exercise 4.3. Use de Moivre's theorem to express $\cos 4\theta$ and $\sin 4\theta$ as polynomials in $\cos\theta$ and $\sin\theta$.

Exercise 4.4. Compute $\operatorname{Log}(-\mathrm{i})$, $\operatorname{Log}(1 + \mathrm{i})$, and all values of $(1 + \mathrm{i})^{\mathrm{i}}$, identifying the principal value of the last.

Exercise 4.5. Show that $\sin z = \sin x \cosh y + \mathrm{i}\cos x \sinh y$, and deduce that $|\sin z|^2 = \sin^2 x + \sinh^2 y$. Explain why this shows $\sin z$ is unbounded on $\mathbb{C}$.

Exercise 4.6. Verify the addition law $\mathrm{e}^{z+w} = \mathrm{e}^z \mathrm{e}^w$ directly from the definition $\mathrm{e}^{x+\mathrm{i}y} = \mathrm{e}^x(\cos y + \mathrm{i}\sin y)$ and the real angle-addition formulas.

Chapter 5

Analytic Functions and the Cauchy–Riemann Equations

5.1 Functions, limits, and continuity

A complex function $w = f(z)$ assigns to each z in a domain $D \subseteq \mathbb{C}$ a complex value w. Separating real and imaginary parts, every such function is a pair of real functions of two real variables,

$$f(z) = u(x, y) + \mathrm{i}\, v(x, y), \qquad z = x + \mathrm{i}y.$$

The notions of limit and continuity read exactly as in $\mathbb{R}^2$: $\lim_{z\to z_0} f(z) = L$ means that $|f(z) - L|$ can be made as small as we please by taking $|z - z_0|$ small enough, and f is *continuous* at z_0 if its limit there equals $f(z_0)$. What distinguishes the complex theory from the real two-variable theory is not these definitions but the notion of derivative to which we turn next; it is far more restrictive, and from that single extra demand flows the entire rich structure of the subject.

5.2 The complex derivative

Definition 5.1. f is *differentiable* at z_0 if the limit

$$f'(z_0) = \lim_{h\to 0} \frac{f(z_0 + h) - f(z_0)}{h}$$

exists, where h ranges over *complex* numbers approaching 0.

The phrase "$h \to 0$ over complex numbers" carries the whole weight of the theory. In one real variable h can approach 0 only from the left or the right; in the plane it may approach from any direction or along any spiral, and the definition demands that the difference quotient tend to the *same* limit $f'(z_0)$ regardless. This is a powerful constraint, and it has an explicit analytic form.

5.3 The Cauchy–Riemann equations

Suppose $f = u + \mathrm{i}v$ is differentiable at $z_0 = x_0 + \mathrm{i}y_0$, and compute the limit in Definition 5.1 along two directions. Letting $h = \Delta x$ be real,

$$f'(z_0) = \lim_{\Delta x\to 0} \frac{f(z_0 + \Delta x) - f(z_0)}{\Delta x} = u_x + \mathrm{i}v_x.$$

Letting instead $h = \mathrm{i}\Delta y$ be purely imaginary,

$$f'(z_0) = \lim_{\Delta y \to 0} \frac{f(z_0 + \mathrm{i}\Delta y) - f(z_0)}{\mathrm{i}\Delta y} = \frac{u_y + \mathrm{i}v_y}{\mathrm{i}} = v_y - \mathrm{i}u_y.$$

Since the two expressions must be equal, their real and imaginary parts match, giving the *Cauchy–Riemann equations*

$$u_x = v_y, \qquad u_y = -v_x. \tag{5.1}$$

Theorem 5.2. *If $f = u + \mathrm{i}v$ is differentiable at z_0, then the partial derivatives of u and v exist there and satisfy the Cauchy–Riemann equations (5.1); moreover $f'(z_0) = u_x + \mathrm{i}v_x$. Conversely, if u and v have continuous first partial derivatives satisfying (5.1) in a neighborhood of z_0, then f is differentiable at z_0.*

The forward direction was just proved; the converse, that continuity of the partials plus (5.1) suffices, is established by writing the increment of f using the differentiability of u and v as real functions and collecting terms with the help of (5.1).

Example 5.3. For $f(z) = \bar{z}$ we have $u = x$, $v = -y$, so $u_x = 1$ but $v_y = -1$: the Cauchy–Riemann equations fail at every point, and $\bar{z}$ is differentiable *nowhere*. For $f(z) = z^2 = (x^2 - y^2) + \mathrm{i}\,2xy$ we have $u_x = 2x = v_y$ and $u_y = -2y = -v_x$, so the equations hold everywhere and z^2 is differentiable on all of $\mathbb{C}$, with $f'(z) = 2x + \mathrm{i}\,2y = 2z$ as expected.

5.4 Analytic and entire functions

Definition 5.4. f is *analytic* (or *holomorphic*) on an open set D if it is differentiable at every point of D. A function analytic on all of $\mathbb{C}$ is *entire*.

Analyticity is a property on an open set, not at a single point—it is the demand that f be differentiable throughout a neighborhood. Polynomials, e^z, $\sin z$, and $\cos z$ are entire; a rational function is analytic except at the zeros of its denominator; and $\bar{z}$, $|z|^2$, and $\operatorname{Re} z$ are nowhere analytic (each fails (5.1) except possibly at isolated points, which are not open sets). The rules of differentiation—sum, product, quotient, and chain—hold with the same statements and the same proofs as in the real case, since those proofs use only the algebra of limits of difference quotients. In particular $\dfrac{d}{dz} z^n = n z^{n-1}$ and $\dfrac{d}{dz}\mathrm{e}^z = \mathrm{e}^z$.

5.5 Harmonic functions and harmonic conjugates

The Cauchy–Riemann equations link analytic functions to the most important partial differential equation of the subject. Differentiating the first equation in (5.1) with respect to x and the second with respect to y, and adding (using equality of mixed partials), gives $u_{xx} + u_{yy} = 0$; the same argument applied the other way gives $v_{xx} + v_{yy} = 0$. Thus the real and imaginary parts of an analytic function are *harmonic*: each satisfies *Laplace's equation*

$$\nabla^2 \phi = \phi_{xx} + \phi_{yy} = 0.$$

We call v a *harmonic conjugate* of u. Given a harmonic u, one constructs v by integrating the Cauchy–Riemann equations: $v_y = u_x$ fixes v up to a function of x, and $v_x = -u_y$ then determines that function up to a constant.

Example 5.5. Let $u = x^2 - y^2$, which is harmonic. From $v_y = u_x = 2x$ we get $v = 2xy + g(x)$, and then $v_x = 2y + g'(x)$ must equal $-u_y = 2y$, forcing $g' \equiv 0$. Hence $v = 2xy + C$, and $f = u + \mathrm{i}v = z^2 + \mathrm{i}C$: we have recovered an analytic function from its real part. This procedure—reconstructing an analytic, hence harmonic, potential from partial data—is the mathematical core of two-dimensional electrostatics and ideal fluid flow.

5.6 Singularities

A point z_0 at which f fails to be analytic, while being analytic at points arbitrarily close to z_0, is a *singularity*. If f is analytic in a punctured disk $0 < |z - z_0| < r$ the singularity is *isolated*. The most important isolated singularities are *poles*: z_0 is a pole of order m if $f(z) = g(z)/(z - z_0)^m$ near z_0 for some analytic g with $g(z_0) \neq 0$, a pole of order one being called *simple*. Singularities that are not poles include *essential singularities*, such as that of $\mathrm{e}^{1/z}$ at the origin, and the *branch points* of multivalued functions like $\log z$ and $\sqrt{z}$. The systematic classification of isolated singularities, by means of the Laurent series, is the subject of the next chapter.

Exercises

Exercise 5.1. For each of $f(z) = z^3$, $f(z) = \operatorname{Re} z$, and $f(z) = z\,\bar{z}$, determine the set of points at which f is differentiable, and state where (if anywhere) f is analytic.

Exercise 5.2. Verify that $u(x, y) = \mathrm{e}^x \cos y$ is harmonic, find a harmonic conjugate v, and identify the resulting analytic function $f = u + \mathrm{i}v$.

Exercise 5.3. Show that if f is analytic on a domain and $f'(z) \equiv 0$ there, then f is constant. (Use the Cauchy–Riemann equations.)

Exercise 5.4. Show that an analytic function with constant modulus $|f|$ on a domain must be constant. (Differentiate $|f|^2 = u^2 + v^2$ and use (5.1).)

Exercise 5.5. *Writing $z = r\mathrm{e}^{\mathrm{i}\theta}$, derive the polar form of the Cauchy–Riemann equations, $u_r = \frac{1}{r} v_\theta$ and $v_r = -\frac{1}{r} u_\theta$, and use them to verify that $\log z$ is analytic away from its branch cut.

Exercise 5.6. Determine the order of the pole of $f(z) = \dfrac{\mathrm{e}^z}{(z^2+1)(z-2)^3}$ at each of its singularities $z = \pm\mathrm{i}$ and $z = 2$.

Chapter 6

Complex Integration

6.1 Contour integrals

Integration in the complex plane is integration along a path. A *contour* C is a piecewise-smooth curve, described by $z(t)$ for $t \in [a, b]$. The *contour integral* of a continuous function f along C is defined by pulling back to the parameter,

$$\int_C f(z)\,dz := \int_a^b f\big(z(t)\big)\,z'(t)\,dt. \tag{6.1}$$

It is linear in f, reverses sign when C is traversed backwards ($\int_{-C} = -\int_C$), and adds over a contour split into pieces. A single estimate, used constantly, controls its size.

Lemma 6.1 (ML-inequality)**.** *If $|f(z)| \le M$ for all z on C, and C has length L, then $\left|\int_C f(z)\,dz\right| \le ML$.*

Proof. Bound the integral (6.1) by $\int_a^b |f(z(t))|\,|z'(t)|\,dt \le M \int_a^b |z'(t)|\,dt = ML$, since $\int_a^b |z'(t)|\,dt$ is the arc length of C. □

6.2 The fundamental example

Almost everything in this chapter grows from a single computation. Let C be the circle $|z - a| = \rho$ traversed once counterclockwise, parametrized by $z = a + \rho e^{i\theta}$, $\theta \in [0, 2\pi]$, so that $dz = i\rho e^{i\theta}\,d\theta$. For an integer n,

$$\oint_C (z-a)^n\,dz = \int_0^{2\pi} \rho^n e^{in\theta}\, i\rho e^{i\theta}\,d\theta = i\rho^{n+1}\int_0^{2\pi} e^{i(n+1)\theta}\,d\theta.$$

The remaining integral is 2π if $n + 1 = 0$ and zero otherwise, because $\int_0^{2\pi} e^{im\theta}\,d\theta = 0$ for every nonzero integer m. Hence

$$\oint_{|z-a|=\rho} (z-a)^n\,dz = \begin{cases} 2\pi i, & n = -1, \\ 0, & n \ne -1. \end{cases} \tag{6.2}$$

The case $n = -1$, $\oint dz/(z-a) = 2\pi i$, is the seed of the entire residue calculus. Note also, by contrast, that $\oint_{|z|=1} \bar{z}\,dz = 2\pi i \ne 0$ even though the contour is closed—a reflection of the fact that $\bar{z}$ is not analytic.

6.3 Cauchy's theorem

A domain is *simply connected* if it has no holes—every closed contour in it can be shrunk to a point without leaving the domain. On such domains the integral of an analytic function around any loop vanishes.

Theorem 6.2 (Cauchy–Goursat)**.** *If f is analytic throughout a simply connected domain D, then $\oint_C f(z)\,dz = 0$ for every closed contour C lying in D.*

Proof (assuming f' continuous). Write $f = u + \mathrm{i}v$ and $dz = dx + \mathrm{i}\,dy$, so that

$$\oint_C f\,dz = \oint_C (u\,dx - v\,dy) + \mathrm{i}\oint_C (v\,dx + u\,dy).$$

By Green's theorem each line integral becomes a double integral over the region R enclosed by C:

$$\oint_C f\,dz = \iint_R (-v_x - u_y)\,dA + \mathrm{i}\iint_R (u_x - v_y)\,dA.$$

Both integrands vanish identically by the Cauchy–Riemann equations (5.1), so the integral is zero. □

The hypothesis that f' be continuous, needed for Green's theorem, is in fact unnecessary: Goursat showed the conclusion holds assuming only that f is differentiable, and it is his sharper form that the theorem properly bears.

6.4 Path independence and antiderivatives

Cauchy's theorem has an immediate and useful reformulation. If f is analytic on a simply connected domain and C_1, C_2 are two contours from z_1 to z_2, then C_1 followed by $-C_2$ is a closed contour, so by Theorem 6.2 the integrals along C_1 and C_2 agree: *the contour integral depends only on the endpoints*. Consequently f possesses an *antiderivative* F (analytic, with $F' = f$), and integration reduces to evaluating it,

$$\int_C f(z)\,dz = F(z_2) - F(z_1).$$

For instance $\int_C z^2\,dz = \frac{1}{3}(z_2^3 - z_1^3)$ and $\int_C \mathrm{e}^z\,dz = \mathrm{e}^{z_2} - \mathrm{e}^{z_1}$ along *any* contour between the endpoints, exactly as in elementary calculus.

6.5 Deformation of contours

When the domain has holes—where f has singularities—loops need not shrink to a point and integrals around them need not vanish. What remains true is the *principle of deformation*: a closed contour may be continuously deformed, within the region where f is analytic, without changing the value of the integral. The reason is again Cauchy's theorem, applied to the region between the original contour and the deformed one.

The deformation principle upgrades the fundamental example (6.2) from circles to arbitrary loops: if C is *any* simple closed contour enclosing the point a once counterclockwise, then shrinking C down to a small circle about a (through the region where $(z-a)^{-1}$ is analytic) shows

$$\oint_C \frac{dz}{z-a} = 2\pi\mathrm{i}.$$

This is the form in which the result is used below.

6.6 Cauchy's integral formula

The values of an analytic function inside a contour are completely determined by its values *on* the contour, as the following theorem states.

Theorem 6.3 (Cauchy's integral formula)**.** *Let f be analytic on and inside a simple closed contour C, traversed counterclockwise, and let a be a point inside C. Then*

$$f(a) = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{f(z)}{z-a}\, dz. \tag{6.3}$$

Proof. Split the integrand:

$$\oint_C \frac{f(z)}{z-a}\, dz = \oint_C \frac{f(z)-f(a)}{z-a}\, dz + f(a) \oint_C \frac{dz}{z-a}.$$

The second integral equals $2\pi\mathrm{i}$ by the previous section. For the first, deform C to a circle of radius ε about a; since f is differentiable at a, the quotient $[f(z)-f(a)]/(z-a)$ stays bounded as $z \to a$, so by the ML-inequality the integral over the small circle is at most (a constant)$\times 2\pi\varepsilon \to 0$ as $\varepsilon \to 0$. Hence the first integral vanishes, and (6.3) follows. □

Differentiating (6.3) under the integral sign with respect to the parameter a yields formulas for all the derivatives,

$$f^{(n)}(a) = \frac{n!}{2\pi \mathrm{i}} \oint_C \frac{f(z)}{(z-a)^{n+1}}\, dz, \qquad n = 0, 1, 2, \ldots \tag{6.4}$$

The consequence is stunning and has no real-variable counterpart: *a function differentiable once on an open set is automatically differentiable infinitely many times there*, since the right-hand side of (6.4) exists for every n. Analyticity, defined as the existence of a single derivative, secretly carries all the others—and, as we shall see in the next chapter, the convergent Taylor series as well.

6.7 Cauchy's inequality and Liouville's theorem

Applying the ML-inequality to the derivative formula (6.4) on a circle of radius R about a, on which $|f| \le M$, gives *Cauchy's inequality*

$$\left| f^{(n)}(a) \right| \le \frac{n!\, M}{R^n}.$$

A single spectacular corollary follows by taking $n = 1$ and letting R grow.

Theorem 6.4 (Liouville)**.** *A bounded entire function is constant.*

Proof. If $|f| \le M$ on all of $\mathbb{C}$, then for every a and every R, Cauchy's inequality gives $|f'(a)| \le M/R$. Letting $R \to \infty$ forces $f'(a) = 0$; since a was arbitrary, $f' \equiv 0$ and f is constant. □

Corollary 6.5 (Fundamental theorem of algebra)**.** ⋆*Every non-constant polynomial $p(z)$ has a root in $\mathbb{C}$.*

Proof. If p had no root, then $1/p$ would be entire; and since $|p(z)| \to \infty$ as $|z| \to \infty$, the function $1/p$ would be bounded. By Liouville's theorem $1/p$, hence p, would be constant, contradicting the hypothesis. □

Exercises

Exercise 6.1. Evaluate $\oint_{|z|=2} \frac{dz}{z-1}$ and $\oint_{|z|=2} \frac{dz}{z-3}$, explaining the difference between the two results in terms of whether the singularity lies inside the contour.

Exercise 6.2. Compute $\int_C \bar{z}\, dz$ where C is (a) the straight segment from 0 to $1+\mathrm{i}$, and (b) the path from 0 to 1 then 1 to $1+\mathrm{i}$. Confirm that the answers differ, and explain why this does not contradict Cauchy's theorem.

Exercise 6.3. Use Cauchy's integral formula to evaluate $\oint_{|z|=1} \frac{\mathrm{e}^z}{z}\, dz$ and $\oint_{|z|=2} \frac{\cos z}{z-\mathrm{i}}\, dz$.

Exercise 6.4. Use the derivative formula (6.4) to evaluate $\oint_{|z|=1} \frac{\mathrm{e}^z}{z^3}\, dz$.

Exercise 6.5. Apply the ML-inequality to show that $\left|\oint_{|z|=R} \frac{dz}{z^2+1}\right| \to 0$ as $R \to \infty$.

Exercise 6.6. *Prove that if f is entire and satisfies $|f(z)| \le A + B\,|z|$ for constants A, B, then f is a polynomial of degree at most one. (Bound f'' using Cauchy's inequality.)

Chapter 7

Taylor and Laurent Series; Residues

7.1 Power series and the radius of convergence

A *power series* about z_0 is a series of the form $\sum_{n=0}^{\infty} a_n(z - z_0)^n$. Its convergence is governed by a single number. By comparison with a geometric series one shows that there is a *radius of convergence* $R \in [0, \infty]$, given by the Cauchy–Hadamard formula $1/R = \limsup_n |a_n|^{1/n}$ (or by the ratio $\lim |a_n/a_{n+1}|$ when that limit exists), such that the series converges absolutely for $|z - z_0| < R$ and diverges for $|z - z_0| > R$. The set $|z - z_0| < R$ is the *disk of convergence*. Inside it the series defines an analytic function, which may be differentiated and integrated term by term.

7.2 Taylor series

The converse—that every analytic function is, locally, a power series—is one of the central theorems of the subject and a direct dividend of the Cauchy integral formula.

Theorem 7.1 (Taylor)**.** *If f is analytic in the disk $|z - z_0| < R$, then for all z in that disk*

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!} (z - z_0)^n, \tag{7.1}$$

and the radius of convergence is at least R. In fact the radius equals the distance from z_0 to the nearest singularity of f.

Proof. Fix z with $|z - z_0| < R$ and apply the Cauchy integral formula (6.3) over a circle C of radius between $|z - z_0|$ and R:

$$f(z) = \frac{1}{2\pi \mathrm{i}} \oint_C \frac{f(w)}{w - z}\, dw.$$

Expand the kernel in a geometric series, valid because $|(z - z_0)/(w - z_0)| < 1$ on C:

$$\frac{1}{w - z} = \frac{1}{(w - z_0) - (z - z_0)} = \sum_{n=0}^{\infty} \frac{(z - z_0)^n}{(w - z_0)^{n+1}}.$$

Insert and integrate term by term; by the derivative formula (6.4) the n-th coefficient is exactly $f^{(n)}(z_0)/n!$, giving (7.1). □

Example 7.2. The geometric and exponential series, $\frac{1}{1-z} = \sum z^n$ (radius 1, the singularity sitting at $z = 1$) and $e^z = \sum z^n/n!$ (radius ∞, e^z entire), are Taylor series. More instructive is $\frac{1}{1+z^2} = \sum(-1)^n z^{2n}$, whose radius of convergence is 1. On the real line this looks mysterious—$1/(1+x^2)$ is perfectly smooth everywhere, yet its Taylor series diverges for $|x| > 1$. The complex picture removes the mystery: $1/(1+z^2)$ has singularities at $z = \pm i$, a distance 1 from the origin, and by Theorem 7.1 no power series about 0 can converge past them.

7.3 Laurent series

Near an isolated singularity a function is not analytic in a full disk, so it has no Taylor series there. It does, however, have a series in an *annulus*—one that allows negative powers.

Theorem 7.3 (Laurent)**.** *If f is analytic in the annulus $r < |z - z_0| < R$, then there*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z - z_0)^n, \qquad c_n = \frac{1}{2\pi i} \oint_C \frac{f(w)}{(w - z_0)^{n+1}}\, dw, \tag{7.2}$$

where C is any circle in the annulus.

The terms with $n \geq 0$ form the *regular part* (convergent inside the outer circle), and the terms with $n < 0$ the *principal part* (convergent outside the inner circle). The proof mirrors Theorem 7.1 but uses two circles bounding the annulus: on the outer one the kernel is expanded as before, while on the inner one the roles of $w - z_0$ and $z - z_0$ are reversed, producing the negative powers.

7.4 Classification of isolated singularities

The principal part of the Laurent series labels the singularity at z_0 precisely:

- *Removable* singularity — no principal part ($c_n = 0$ for all $n < 0$). Then f extends analytically across z_0. Example: $\sin z/z = 1 - z^2/3! + \cdots$ near 0.
- *Pole of order m* — finitely many negative terms, the lowest being $c_{-m}(z - z_0)^{-m}$ with $c_{-m} \neq 0$. Example: $1/z^m$.
- *Essential* singularity — infinitely many negative terms. Example: $e^{1/z} = \sum_{n \geq 0} z^{-n}/n!$ at 0.

Remark 7.4. *Essential singularities are genuinely wild: by the Casorati–Weierstrass theorem, in every neighborhood of an essential singularity f comes arbitrarily close to *every* complex value, and Picard's great theorem sharpens this to say f actually attains every value, with at most one exception, infinitely often.

7.5 Residues

Definition 7.5. The *residue* of f at an isolated singularity z_0 is the Laurent coefficient

$$\operatorname{Res}(f, z_0) := c_{-1}.$$

The residue is singled out by the keystone integral (6.2): of all the powers $(z - z_0)^n$ in the Laurent series, only $(z - z_0)^{-1}$ has a nonzero integral around z_0. Integrating (7.2) term by term over a small circle therefore returns $2\pi i\, c_{-1}$, a fact we elevate to a theorem in the next section. For the common case of poles, the residue is computed without finding the whole series.

Proposition 7.6. *If f has a simple pole at z_0, then* $\operatorname{Res}(f, z_0) = \lim_{z\to z_0}(z - z_0)f(z)$*. If f has a pole of order m, then*

$$\operatorname{Res}(f, z_0) = \frac{1}{(m-1)!} \lim_{z\to z_0} \frac{d^{m-1}}{dz^{m-1}}\big[(z - z_0)^m f(z)\big].$$

In particular, if $f = p/q$ with p, q analytic, $q(z_0) = 0$, $q'(z_0) \neq 0$, and $p(z_0) \neq 0$, then the pole is simple and $\operatorname{Res}(f, z_0) = p(z_0)/q'(z_0)$*.*

Proof. For a pole of order m, $(z - z_0)^m f(z)$ is analytic at z_0 with Taylor series $\sum_{k\geq 0} c_{k-m}(z - z_0)^k$; the coefficient c_{-1} is the one of $(z - z_0)^{m-1}$, extracted by differentiating $m - 1$ times and dividing by $(m - 1)!$. The quotient formula is the simple-pole case with $\lim(z - z_0)p/q = p(z_0)/q'(z_0)$ by l'Hôpital. □

7.6 The Residue Theorem

Theorem 7.7 (Residue Theorem)**.** *Let f be analytic on and inside a simple closed contour C (counterclockwise) except for isolated singularities $z_1, \ldots, z_N$ inside C. Then*

$$\oint_C f(z)\,dz = 2\pi \mathrm{i} \sum_{k=1}^{N} \operatorname{Res}(f, z_k). \tag{7.3}$$

Proof. Surround each z_k by a small circle γ_k lying inside C and enclosing no other singularity. In the region between C and the γ_k, f is analytic, so by contour deformation $\oint_C f\,dz = \sum_k \oint_{\gamma_k} f\,dz$. On γ_k expand f in its Laurent series about z_k and integrate term by term; by (6.2) every term integrates to zero except $c_{-1}^{(k)}(z - z_k)^{-1}$, which contributes $2\pi\mathrm{i}\, c_{-1}^{(k)} = 2\pi\mathrm{i} \operatorname{Res}(f, z_k)$. Summing over k gives (7.3). □

The Residue Theorem is the grand unification of the chapter: Cauchy's theorem is its special case with no singularities (empty sum, integral zero), and the Cauchy integral formula is the case of a single simple pole. It reduces the computation of a contour integral to the purely local, algebraic task of finding residues.

Example 7.8. To evaluate $\oint_{|z|=2} \dfrac{z}{(z-1)(z+3)}\,dz$, note that only the simple pole $z = 1$ lies inside the contour ($z = -3$ is outside). By the quotient formula its residue is $\dfrac{1}{1+3} = \frac{1}{4}$, so the integral is $2\pi\mathrm{i} \cdot \frac{1}{4} = \frac{\pi \mathrm{i}}{2}$.

Exercises

Exercise 7.1. Find the Taylor series of $f(z) = 1/(2 - z)$ about $z_0 = 0$ and about $z_0 = 1$, and give the radius of convergence in each case, identifying the singularity that limits it.

Exercise 7.2. Find the Laurent series of $f(z) = \dfrac{1}{z(z-1)}$ valid in (a) the annulus $0 < |z| < 1$, and (b) the region $|z| > 1$. (Hint: partial fractions.)

Exercise 7.3. Classify the singularity at $z = 0$ of each of $\dfrac{1 - \cos z}{z^2}$, $\dfrac{\sin z}{z^3}$, and $z\,\mathrm{e}^{1/z}$, and give the residue there.

Exercise 7.4. Compute $\operatorname{Res}\left(\frac{e^z}{z^2(z-1)}, 0\right)$ and $\operatorname{Res}\left(\frac{e^z}{z^2(z-1)}, 1\right)$.

Exercise 7.5. Use the Residue Theorem to evaluate $\oint_{|z|=3} \frac{dz}{z^2+1}$ and $\oint_{|z|=3} \frac{z\,dz}{(z-1)^2(z-i)}$.

Exercise 7.6. *Show that if g is analytic and nonzero at z_0, the residue of $g(z)/(z-z_0)^m$ at z_0 equals $g^{(m-1)}(z_0)/(m-1)!$, and use this to recover Proposition 7.6.

Chapter 8

Evaluation of Real Integrals by Residues

8.1 The method

The Residue Theorem furnishes a powerful and often startling technique for evaluating definite *real* integrals that resist elementary methods. The strategy is to regard the real integral as part of a closed contour integral in the complex plane: one completes the path of integration with an auxiliary arc, evaluates the closed integral by residues, and shows that the contribution of the arc vanishes (or is otherwise computable). What is left is the real integral in terms of residues. We organize the chapter by the type of integrand.

8.2 Rational functions over the line

Consider $\displaystyle\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)}\,dx$ where P, Q are polynomials, Q has no real zeros, and $\deg Q \geq \deg P + 2$ (so the integral converges). Close the interval $[-R, R]$ with the semicircular arc C_R of radius R in the upper half-plane. On C_R the integrand is $O(R^{-2})$ while the arc has length πR, so by the ML-inequality its contribution is $O(R^{-1}) \to 0$. In the limit the Residue Theorem gives

$$\int_{-\infty}^{\infty} \frac{P(x)}{Q(x)}\,dx = 2\pi\mathrm{i} \sum_{\operatorname{Im} z_k > 0} \operatorname{Res}\left(\frac{P}{Q}, z_k\right), \tag{8.1}$$

the sum over poles in the upper half-plane.

Example 8.1. For $\int_{-\infty}^{\infty} dx/(1+x^2)$, the only pole in the upper half-plane is $z = \mathrm{i}$, with residue $1/(2\mathrm{i})$, so the integral is $2\pi\mathrm{i} \cdot \frac{1}{2\mathrm{i}} = \pi$. For $\int_{-\infty}^{\infty} dx/(1+x^4)$, the upper-half poles are $\mathrm{e}^{\mathrm{i}\pi/4}$ and $\mathrm{e}^{\mathrm{i}3\pi/4}$; using $\operatorname{Res}(1/(z^4+1), z_k) = 1/(4z_k^3) = -z_k/4$ and summing gives $2\pi\mathrm{i}\left(-\frac{1}{4}\right)(\mathrm{e}^{\mathrm{i}\pi/4} + \mathrm{e}^{\mathrm{i}3\pi/4}) = \pi/\sqrt{2}$.

8.3 Fourier-type integrals and Jordan's lemma

Integrals of the form $\int_{-\infty}^{\infty} F(x)\mathrm{e}^{\mathrm{i}mx}\,dx$ with $m > 0$ are handled the same way, closing in the upper half-plane, where $\left|\mathrm{e}^{\mathrm{i}mz}\right| = \mathrm{e}^{-m\operatorname{Im} z} \leq 1$ decays. The decay is enough to kill the arc even when F falls off only like $1/R$, a fact made precise by *Jordan's lemma*: if $F(z) \to 0$ uniformly on the upper semicircle as $R \to \infty$, then $\int_{C_R} F(z)\mathrm{e}^{\mathrm{i}mz}\,dz \to 0$ for every $m > 0$. Hence

$$\int_{-\infty}^{\infty} F(x)\mathrm{e}^{\mathrm{i}mx}\,dx = 2\pi\mathrm{i} \sum_{\operatorname{Im} z_k > 0} \operatorname{Res}\big(F(z)\mathrm{e}^{\mathrm{i}mz}, z_k\big),$$

and real integrals of $F(x)\cos mx$ or $F(x)\sin mx$ are recovered as the real or imaginary part.

Example 8.2. $\displaystyle\int_{-\infty}^{\infty} \frac{\cos mx}{x^2+1}\,dx$ is the real part of $\int \mathrm{e}^{\mathrm{i}mx}/(x^2+1)\,dx = 2\pi\mathrm{i}\ \operatorname{Res}_{z=\mathrm{i}} = 2\pi\mathrm{i}\cdot\dfrac{\mathrm{e}^{-m}}{2\mathrm{i}} = \pi\mathrm{e}^{-m}$. Thus the integral equals $\pi\mathrm{e}^{-m}$ for $m > 0$.

8.4 Trigonometric integrals over a period

An integral $\int_0^{2\pi} R(\cos\theta, \sin\theta)\,d\theta$ of a rational function of $\cos\theta$ and $\sin\theta$ becomes a contour integral around the unit circle under the substitution $z = \mathrm{e}^{\mathrm{i}\theta}$, for which

$$\cos\theta = \frac{z + z^{-1}}{2}, \qquad \sin\theta = \frac{z - z^{-1}}{2\mathrm{i}}, \qquad d\theta = \frac{dz}{\mathrm{i}z}.$$

The integral turns into $\oint_{|z|=1}$ of a rational function of z, evaluated by the residues inside the unit circle.

Example 8.3. For $a > |b| > 0$,

$$\int_0^{2\pi} \frac{d\theta}{a + b\cos\theta} = \oint_{|z|=1} \frac{2\,dz}{\mathrm{i}\,(bz^2 + 2az + b)}.$$

The denominator's two roots multiply to 1, so exactly one, z_+, lies inside the unit circle; computing its residue gives the value $\dfrac{2\pi}{\sqrt{a^2 - b^2}}$.

8.5 Integrals with branch cuts

Integrals over the half-line involving a fractional power require a contour that respects the branch cut of that power. The standard device is the *keyhole contour*, which runs out along the top of the positive real axis, around a large circle, back along the bottom of the axis, and around a small circle at the origin, with the branch cut of z^{a-1} taken along the positive real axis so that $\arg z$ runs from 0 to 2π.

Example 8.4. ⋆For $0 < a < 1$, the keyhole contour applied to $z^{a-1}/(1+z)$ yields

$$\int_0^{\infty} \frac{x^{a-1}}{1+x}\,dx = \frac{\pi}{\sin\pi a}.$$

The two straight portions of the contour differ by the factor $\mathrm{e}^{2\pi\mathrm{i}(a-1)}$ picked up by z^{a-1} in circling the origin, while the two circular portions vanish in the limits; the single pole at $z = -1$ supplies the residue, and rearranging gives the stated value.

8.6 Indented contours and principal values

When the integrand has a singularity *on* the contour of integration, one indents the path with a small semicircle around the singular point and interprets the integral as a Cauchy principal value. A small semicircle of radius ε about a simple pole contributes $\pm\pi\mathrm{i}$ times the residue (half of the full $2\pi\mathrm{i}$, with sign set by orientation) as $\varepsilon \to 0$.

Example 8.5. ⋆The classic $\displaystyle\int_0^\infty \frac{\sin x}{x}\,dx = \frac{\pi}{2}$ is obtained by integrating $\mathrm{e}^{\mathrm{i}z}/z$ around a contour indented at the origin: the large arc vanishes by Jordan's lemma, the small indentation contributes $-\pi\mathrm{i}\ \mathrm{Res}_{z=0} = -\pi\mathrm{i}$, and taking the imaginary part of the resulting identity isolates the integral.

Exercises

Exercise 8.1. Evaluate $\displaystyle\int_{-\infty}^\infty \frac{dx}{(x^2+1)^2}$ by closing in the upper half-plane. (The pole at $z = \mathrm{i}$ is now of order two.)

Exercise 8.2. Evaluate $\displaystyle\int_{-\infty}^\infty \frac{x^2}{x^4+1}\,dx$ using (8.1).

Exercise 8.3. Show that $\displaystyle\int_{-\infty}^\infty \frac{\cos x}{x^2+4}\,dx = \frac{\pi}{2}\mathrm{e}^{-2}$ via Jordan's lemma.

Exercise 8.4. Evaluate $\displaystyle\int_0^{2\pi} \frac{d\theta}{5+4\cos\theta}$ by the unit-circle substitution.

Exercise 8.5. Use the substitution $z = \mathrm{e}^{\mathrm{i}\theta}$ to show $\displaystyle\int_0^{2\pi} \cos^{2n}\theta\,d\theta = 2\pi\binom{2n}{n}4^{-n}$.

Exercise 8.6. ⋆Carry out the keyhole-contour evaluation of $\int_0^\infty x^{a-1}/(1+x)\,dx$ in detail for $0 < a < 1$, justifying the vanishing of the two circular arcs.

Chapter 9

Conformal Mapping and Its Applications

9.1 Conformal maps

An analytic function does more than transform points: where its derivative is nonzero it transforms *shapes*, and it does so in a remarkably rigid way.

Theorem 9.1. *If f is analytic at z_0 and $f'(z_0) \neq 0$, then f is* conformal *at z_0: it preserves the angle between any two curves through z_0, in both magnitude and orientation.*

Proof. Let a smooth curve through z_0 be $z(t)$ with $z(0) = z_0$ and tangent $z'(0) \neq 0$. Its image $f(z(t))$ has tangent $\big(f \circ z\big)'(0) = f'(z_0)z'(0)$. Thus every tangent vector at z_0 is multiplied by the same complex number $f'(z_0)$, which rotates it by $\arg f'(z_0)$ and scales it by $|f'(z_0)|$. Two curves are therefore turned by the same angle, so the angle between them is unchanged. □

A conformal map is thus, infinitesimally, a rotation together with a uniform scaling—it preserves angles but not, in general, lengths or shapes in the large. The points where $f' = 0$ are exceptional, and angles there are typically multiplied.

9.2 Elementary maps and Mobius transformations

The simplest building blocks are translation $z \mapsto z + b$, rotation and scaling $z \mapsto az$, and inversion $z \mapsto 1/z$. Each carries the family of "circles and lines" into itself—a line being regarded as a circle through the point at infinity. Composing these yields the most important family of maps.

Definition 9.2. A *Mobius* (or linear fractional) transformation is

$$w = \frac{az + b}{cz + d}, \qquad ad - bc \neq 0.$$

Mobius transformations form a group under composition, are bijections of the extended plane $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, and map circles-and-lines to circles-and-lines. A Mobius transformation is uniquely determined by its values at any three distinct points, which makes them ideally suited to mapping one specified region onto another. More generally, the *Riemann mapping theorem* guarantees that any simply connected domain other than the whole plane can be mapped conformally onto the unit disk; the practical art is to find an explicit map for the domain at hand.

9.3 Conformal mapping and Laplace's equation

The reason conformal maps are central to physics is that they preserve the harmonicity that governs steady states.

Theorem 9.3. *Let $w = f(z)$ be conformal, and let $\Phi(u, v)$ be harmonic. Then the pulled-back function $\Psi(x, y) = \Phi\big(u(x, y), v(x, y)\big)$ is harmonic.*

A short computation with the chain rule and the Cauchy–Riemann equations gives the clean identity $\Psi_{xx} + \Psi_{yy} = |f'(z)|^2\,(\Phi_{uu} + \Phi_{vv})$, from which the theorem is immediate. This furnishes a general method for boundary value problems: *to solve Laplace's equation on a complicated domain D, find a conformal map f taking D to a simple domain D' (a half-plane or a disk), solve the transformed problem on D', and pull the solution back by $\Psi = \Phi \circ f$.* The boundary data are carried along by the map.

Example 9.4 (A boundary value problem)**.** Solve $\nabla^2\Phi = 0$ in the infinite strip $0 < y < a$ with $\Phi = \alpha$ on the edge $y = 0$ and $\Phi = \beta$ on the edge $y = a$. The map $w = \mathrm{e}^{\pi z/a}$ sends the strip conformally onto the upper half-plane, carrying the edge $y = 0$ to the positive real axis ($\arg w = 0$) and the edge $y = a$ to the negative real axis ($\arg w = \pi$). In the half-plane the harmonic function taking the value α on the positive axis and β on the negative axis is, by inspection (or by Poisson's formula (3.4)), $\Phi = \alpha + (\beta - \alpha)\dfrac{\arg w}{\pi}$. Since $\arg w = \arg \mathrm{e}^{\pi z/a} = \pi y/a$, pulling back gives

$$\Phi(x, y) = \alpha + (\beta - \alpha)\frac{y}{a},$$

the expected linear temperature profile across the strip—now derived by transplanting a trivial problem from the half-plane.

9.4 Ideal fluid flow

The same complex-analytic machinery describes the steady flow of an ideal fluid. Consider a flow in the plane with velocity field $\mathbf{V} = (u, v)$ that is *incompressible* ($\nabla \cdot \mathbf{V} = 0$) and *irrotational* ($\nabla \times \mathbf{V} = \mathbf{0}$). Irrotationality gives a *velocity potential* ϕ with $\mathbf{V} = \nabla\phi$, and incompressibility then makes ϕ harmonic, $\nabla^2\phi = 0$. Let ψ be a harmonic conjugate of ϕ; the analytic combination

$$\Omega(z) = \phi(x, y) + \mathrm{i}\,\psi(x, y)$$

is the *complex potential*, and ψ is the *stream function*. Differentiating and using the Cauchy–Riemann equations,

$$\Omega'(z) = \phi_x + \mathrm{i}\psi_x = \phi_x - \mathrm{i}\phi_y = u - \mathrm{i}v,$$

so the velocity is recovered as $u + \mathrm{i}v = \overline{\Omega'(z)}$. The level curves $\psi = \text{const}$ are the *streamlines* of the flow—the paths the fluid follows—because $\mathbf{V} = \nabla\phi$ is orthogonal to $\nabla\psi$. A solid boundary, which the fluid cannot cross, must itself be a streamline.

Example 9.5 (Three flows)**.** *Uniform flow:* $\Omega(z) = Uz$ gives $\Omega' = U$, a constant velocity U to the right, with horizontal streamlines $\psi = Uy = \text{const}$. *Flow in a right-angle corner:* $\Omega(z) = Az^2$ (A real) gives stream function $\psi = 2Axy$, whose streamlines $xy = \text{const}$ are hyperbolas hugging the corner formed by the positive axes. *Flow past a cylinder:* $\Omega(z) = U\left(z + \dfrac{a^2}{z}\right)$ models uniform flow past a circular cylinder of radius a: on $|z| = a$ one computes $\psi = 0$, so the circle is a streamline, and far away $\Omega' \to U$ recovers the uniform stream. This single formula captures the classical picture of an ideal fluid streaming around a cylinder.

Exercises

Exercise 9.1. Show that the map $w = 1/z$ sends the vertical line $\operatorname{Re} z = c$ (with $c \neq 0$) to a circle, and find its center and radius.

Exercise 9.2. Find the Mobius transformation that sends $z = 0, 1, \infty$ to $w = 1, \mathrm{i}, -1$ respectively, and verify it maps the real axis to a circle through those three points.

Exercise 9.3. Using the map $w = \mathrm{e}^{\pi z/a}$, solve Laplace's equation in the strip $0 < y < a$ with $\Phi(x, 0) = 0$ and $\Phi(x, a) = 1$, and confirm the linear profile by direct substitution.

Exercise 9.4. For the complex potential $\Omega(z) = Az^2$, sketch several streamlines and verify that the positive x- and y-axes together form the streamline $\psi = 0$, so that the flow occupies a right-angle corner.

Exercise 9.5. For flow past a cylinder, $\Omega(z) = U(z + a^2/z)$, compute the velocity $\overline{\Omega'(z)}$ and locate the two *stagnation points* (where the velocity vanishes) on the cylinder.

Exercise 9.6. *Prove Theorem 9.3 by verifying the identity $\Psi_{xx} + \Psi_{yy} = |f'(z)|^2 \, (\Phi_{uu} + \Phi_{vv})$ with the chain rule and the Cauchy–Riemann equations for $f = u + \mathrm{i}v$.

Chapter 10

Vector Spaces and Linear Operators

10.1 Vector spaces

The objects of the previous parts—geometric vectors, functions, sequences of Fourier coefficients—share a common algebraic skeleton: they can be added together and scaled by numbers. Abstracting that skeleton gives the single most unifying notion of the course.

Definition 10.1. A *vector space* over a field of scalars (for us $\mathbb{R}$ or $\mathbb{C}$) is a set V with an addition $\mathbf{u}+\mathbf{v}$ and a scalar multiplication $\alpha\mathbf{u}$ satisfying, for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and scalars α, β:

(i) addition is associative and commutative;

(ii) there is a zero vector $\mathbf{0}$ with $\mathbf{u}+\mathbf{0}=\mathbf{u}$, and each $\mathbf{u}$ has a negative $-\mathbf{u}$;

(iii) $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$ and $1\,\mathbf{u} = \mathbf{u}$;

(iv) the distributive laws $\alpha(\mathbf{u}+\mathbf{v}) = \alpha\mathbf{u}+\alpha\mathbf{v}$ and $(\alpha+\beta)\mathbf{u} = \alpha\mathbf{u}+\beta\mathbf{u}$ hold.

From these axioms the expected small facts follow—the zero vector is unique, $0\,\mathbf{u} = \mathbf{0}$, and $(-1)\mathbf{u} = -\mathbf{u}$. The reach of the definition is best seen in its examples: the coordinate spaces $\mathbb{R}^n$ and $\mathbb{C}^n$; the space $C[a,b]$ of continuous functions on an interval, with $(f+g)(x) = f(x)+g(x)$; the polynomials of degree at most n; and—crucially for this course—infinite-dimensional spaces of functions, where the vectors are themselves functions. Everything below applies verbatim to all of them.

10.2 Independence, basis, and dimension

A *subspace* is a subset closed under addition and scalar multiplication. The *span* of a set of vectors is the collection of all their linear combinations. Vectors $\mathbf{v}_1, \dots, \mathbf{v}_k$ are *linearly independent* if $\sum c_i\mathbf{v}_i = \mathbf{0}$ forces every $c_i = 0$, and *dependent* otherwise. A *basis* of V is a linearly independent set that spans V; the number of vectors in a basis is the *dimension* $\dim V$, independent of the chosen basis. The defining virtue of a basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ is that every $\mathbf{u} \in V$ has a *unique* expansion

$$\mathbf{u} = \sum_{i=1}^{n} u_i\,\mathbf{v}_i,$$

whose coefficients u_i are the *components* of $\mathbf{u}$ in that basis. Choosing a basis turns abstract vectors into columns of numbers, and linear maps into matrices, as we now make precise.

10.3 Matrices and matrix algebra

An $m \times n$ matrix is a rectangular array $\mathbf{A} = (a_{ij})$ of scalars. Matrices of a fixed size add entrywise and scale entrywise, and under these operations $\mathrm{Mat}_{m,n}$ is itself a vector space of dimension mn, a basis being the matrices $\mathbf{E}^{(k,l)}$ with a single 1 in position (k,l). The *transpose* $\mathbf{A}^\top$ interchanges rows and columns, and *matrix multiplication* is defined by $(\mathbf{AB})_{ij} = \sum_k a_{ik} b_{kj}$. Multiplication is associative and distributes over addition, satisfies $(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top$, but is *not* commutative in general. The $n \times n$ identity $\mathbf{I} = (\delta_{ij})$ is the multiplicative unit.

10.4 Determinant and trace

Two scalars summarize a square matrix. The *determinant* may be defined through the permutation (Levi-Civita) formula

$$\det \mathbf{A} = \sum_\sigma \mathrm{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \varepsilon_{i_1 \cdots i_n}\, a_{1i_1} \cdots a_{ni_n},$$

the sum over permutations σ of $\{1, \ldots, n\}$. It is multiplicative, $\det(\mathbf{AB}) = \det \mathbf{A}\, \det \mathbf{B}$, unchanged by transposition, and—decisively—$\mathbf{A}$ is invertible if and only if $\det \mathbf{A} \neq 0$. The *trace* is the sum of the diagonal entries, $\mathrm{tr}\, \mathbf{A} = \sum_i a_{ii}$; it is linear and has the cyclic property $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$, from which it is invariant under the change of basis described next.

10.5 Linear operators and their matrices

A *linear operator* on V is a map $T : V \to V$ with $T(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha\, T\mathbf{u} + \beta\, T\mathbf{v}$. Once a basis $\{\mathbf{v}_j\}$ is fixed, T is encoded by the matrix $\mathbf{A} = (a_{ij})$ whose columns are the images of the basis vectors, $T\mathbf{v}_j = \sum_i a_{ij} \mathbf{v}_i$. Operators then mirror matrices exactly: the composition $S \circ T$ corresponds to the product of the matrices (in the same order), and the action $T\mathbf{u}$ corresponds to multiplying the component column by $\mathbf{A}$. Changing to a new basis related by an invertible matrix $\mathbf{S}$ replaces the matrix of T by the *similar* matrix $\mathbf{S}^{-1}\mathbf{AS}$. The quantities $\det \mathbf{A}$ and $\mathrm{tr}\, \mathbf{A}$ are unchanged by this replacement and so are genuine properties of the operator, not of the basis.

Exercises

Exercise 10.1. Show that the set of real 2×2 symmetric matrices is a subspace of $\mathrm{Mat}_{2,2}(\mathbb{R})$, and find its dimension and a basis.

Exercise 10.2. Decide whether the polynomials $1,\ 1 + x,\ 1 + x + x^2$ form a basis of the space of polynomials of degree at most two, and if so express x^2 in this basis.

Exercise 10.3. Verify the cyclic property $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$ from the definition, and deduce that $\mathrm{tr}(\mathbf{S}^{-1}\mathbf{AS}) = \mathrm{tr}\, \mathbf{A}$.

Exercise 10.4. Give an explicit pair of 2×2 matrices with $\mathbf{AB} \neq \mathbf{BA}$, and compute $\det(\mathbf{AB})$ and $\det(\mathbf{BA})$ to confirm they are nonetheless equal.

Exercise 10.5. Let T be the operator "differentiate" on the space of polynomials of degree at most three. Write its matrix in the basis $\{1, x, x^2, x^3\}$ and compute $\det T$ and $\mathrm{tr}\, T$.

Exercise 10.6. *Show that similar matrices have the same characteristic polynomial $\det(\mathbf{A} - \lambda\mathbf{I})$, and hence the same eigenvalues.

Chapter 11

Inner-Product Spaces and Orthogonal Functions

11.1 Norms and inner products

To do geometry in a vector space—to speak of lengths, angles, and orthogonality—we equip it with an inner product.

Definition 11.1. An *inner product* on a complex vector space V assigns to each pair a scalar $\langle \mathbf{u}, \mathbf{v} \rangle$ that is conjugate-linear in the first argument and linear in the second, satisfies $\langle \mathbf{u}, \mathbf{v} \rangle = \overline{\langle \mathbf{v}, \mathbf{u} \rangle}$, and is positive-definite: $\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ with equality only for $\mathbf{u} = \mathbf{0}$. The associated *norm* is $\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$.

Two vectors are *orthogonal* when their inner product vanishes. The model example is the Euclidean inner product $\langle \mathbf{u}, \mathbf{v} \rangle = \sum_i \overline{u_i}\, v_i$ on $\mathbb{C}^n$, but the same axioms support a vast range of other choices.

11.2 The Cauchy–Schwarz inequality

Theorem 11.2 (Cauchy–Schwarz)**.** *For all* $\mathbf{u}, \mathbf{v}$ *in an inner-product space,* $|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\ \|\mathbf{v}\|$*, with equality precisely when* $\mathbf{u}, \mathbf{v}$ *are linearly dependent.*

Proof. The inequality is clear if $\mathbf{v} = \mathbf{0}$; otherwise set $\lambda = \langle \mathbf{v}, \mathbf{u} \rangle / \langle \mathbf{v}, \mathbf{v} \rangle$. Expanding the nonnegative quantity $\langle \mathbf{u} - \lambda \mathbf{v}, \mathbf{u} - \lambda \mathbf{v} \rangle$ and simplifying gives

$$0 \leq \langle \mathbf{u} - \lambda \mathbf{v}, \mathbf{u} - \lambda \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle - \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\langle \mathbf{v}, \mathbf{v} \rangle},$$

which rearranges to the stated inequality. Equality forces $\mathbf{u} = \lambda \mathbf{v}$. □

Cauchy–Schwarz in turn yields the triangle inequality $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$, so a norm deserves its name. Not every norm comes from an inner product, however; on $\mathbb{C}^n$ the family $\|\mathbf{x}\|_p = (\sum_i |x_i|^p)^{1/p}$ for $1 \leq p \leq \infty$ (with $\|\mathbf{x}\|_\infty = \max_i |x_i|$) are norms, and only $p = 2$ arises from an inner product.

11.3 Function spaces and weight functions

The decisive step—the one that powers the entire method of eigenfunction expansions—is to treat *functions* as vectors and integrate to form their inner product. On a space of functions on $[a, b]$ one sets

$$\langle f, g\rangle = \int_a^b \overline{f(x)}\, g(x)\, w(x)\, dx, \tag{11.1}$$

where $w(x) > 0$ is a fixed *weight function.* The resulting norm is the L^2 norm $\|f\| = (\int_a^b |f|^2\, w\, dx)^{1/2}$, and the space of functions with finite norm is the central infinite-dimensional inner-product space $L^2([a, b], w)$. Orthogonality of functions now means $\int_a^b \overline{f} g\, w\, dx = 0$—exactly the relation that made the Fourier coefficients of Chapter 1 computable.

11.4 Orthogonal polynomials

Applying the Gram–Schmidt process to the monomials $1, x, x^2, \ldots$ under an inner product (11.1) produces a sequence of mutually orthogonal polynomials, one of each degree. Different weights and intervals give the classical families, each ubiquitous in physics:

Family	Interval	Weight $w(x)$	Arises in
Legendre	$[-1, 1]$	1	potential theory, the sphere
Chebyshev	$[-1, 1]$	$(1 - x^2)^{-1/2}$	approximation theory
Hermite	$(-\infty, \infty)$	e^{-x^2}	the quantum oscillator
Laguerre	$[0, \infty)$	e^{-x}	the hydrogen atom

The first few Legendre polynomials, for instance, are $P_0 = 1$, $P_1 = x$, $P_2 = \frac{1}{2}(3x^2 - 1)$, and one checks directly that $\int_{-1}^{1} P_m P_n\, dx = 0$ for $m \neq n$. Each family turns out to consist of the eigenfunctions of a self-adjoint (Sturm–Liouville) differential operator, which is the deep reason such different-looking polynomials all share the orthogonality property—a theme developed in courses on boundary value problems.

11.5 Dirac notation and the dual space

Physics employs a notation, due to Dirac, that makes the algebra of inner products especially transparent. A vector is written as a *ket* $|u\rangle$, and to each vector is associated a *bra* $\langle u|$, the linear functional $|v\rangle \mapsto \langle u|v\rangle := \langle \mathbf{u}, \mathbf{v}\rangle$. The bras form the *dual space* V^* of linear functionals on V, and the *Riesz representation theorem* asserts that in an inner-product space the correspondence is exact: every continuous linear functional is $\langle w|$ for a unique vector $|w\rangle$. Juxtaposing a ket and a bra the other way produces an *operator*,

$$\big(|u\rangle\langle v|\big)|x\rangle = \langle v|x\rangle\, |u\rangle,$$

and for a unit vector $|e\rangle$ the operator $|e\rangle\langle e|$ is the orthogonal projection onto the line through $|e\rangle$.

11.6 Orthonormal bases and generalized Fourier series

A basis $\{|e_i\rangle\}$ is *orthonormal* if $\langle e_i|e_j\rangle = \delta_{ij}$. In such a basis every vector is reconstructed by projecting onto each axis,

$$|u\rangle = \sum_i \langle e_i|u\rangle\, |e_i\rangle, \tag{11.2}$$

so the components are simply the inner products $c_i = \langle e_i|u\rangle$. The identity expressed by reading (11.2) as an operator equation,

$$\sum_i |e_i\rangle\langle e_i| = \mathbf{I},$$

is the *completeness relation* (or resolution of the identity), and the norm obeys *Parseval's relation* $\|u\|^2 = \sum_i |c_i|^2$. This is the abstract content of Fourier analysis: the exponentials $|e_n\rangle = \frac{1}{\sqrt{2L}} \mathrm{e}^{\mathrm{i}n\pi x/L}$ form an orthonormal basis of $L^2([-L, L])$, the Fourier coefficient of Chapter 1 is the projection $\langle e_n|f\rangle$, and Parseval's theorem there is Parseval's relation here. A Fourier series is nothing but the expansion (11.2) in a particular orthonormal basis.

Exercises

Exercise 11.1. On $\mathbb{C}^2$ with the Euclidean inner product, verify the Cauchy–Schwarz inequality for $\mathbf{u} = (1, \mathrm{i})$ and $\mathbf{v} = (1, 1)$, and determine the angle (via $\operatorname{Re}\langle \mathbf{u}, \mathbf{v}\rangle$) between their real parts.

Exercise 11.2. Using the inner product (11.1) with $w = 1$ on $[-1, 1]$, apply Gram–Schmidt to $1, x, x^2$ to obtain the first three Legendre polynomials (up to normalization).

Exercise 11.3. Show that the functions $\{\mathrm{e}^{\mathrm{i}n\pi x/L}\}_{n\in\mathbb{Z}}$ are orthogonal on $[-L, L]$ under (11.1) with $w = 1$, and find the normalization that makes them orthonormal.

Exercise 11.4. For a unit vector $|e\rangle$, verify that $P = |e\rangle\langle e|$ satisfies $P^2 = P$ (idempotent) and interpret this geometrically.

Exercise 11.5. Prove the parallelogram law $\|\mathbf{u} + \mathbf{v}\|^2 + \|\mathbf{u} - \mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$, and explain why it can fail for the p-norm when $p \neq 2$.

Exercise 11.6. *Use the completeness relation to derive Parseval's relation $\|u\|^2 = \sum_i |\langle e_i|u\rangle|^2$ for an orthonormal basis.

Chapter 12

Eigenvalues, Hermitian Operators, and the Matrix Exponential

12.1 Eigenvalues and eigenvectors

Definition 12.1. A nonzero vector $\mathbf{v}$ is an *eigenvector* of the operator $\mathbf{A}$ with *eigenvalue* λ if $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$.

An eigenvector is a direction left invariant by $\mathbf{A}$, merely stretched by the factor λ. Since $\mathbf{A}\mathbf{v} = \lambda\mathbf{v}$ means $(\mathbf{A} - \lambda\mathbf{I})\mathbf{v} = \mathbf{0}$ has a nonzero solution, the eigenvalues are exactly the roots of the *characteristic equation*

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

If an $n \times n$ matrix has n distinct eigenvalues, the corresponding eigenvectors are linearly independent and hence form a basis: such a matrix is *diagonalizable*.

12.2 Diagonalization

Collecting eigenvectors as the columns of a matrix $\mathbf{S}$ and the eigenvalues into a diagonal matrix $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, the relations $\mathbf{A}\mathbf{v}_i = \lambda_i\mathbf{v}_i$ combine into $\mathbf{A}\mathbf{S} = \mathbf{S}\mathbf{\Lambda}$, that is

$$\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}.$$

In the eigenbasis the operator is just a list of independent scalings. This representation trivializes functions of $\mathbf{A}$: powers become $\mathbf{A}^k = \mathbf{S}\mathbf{\Lambda}^k\mathbf{S}^{-1}$, and more generally any power series $f(\mathbf{A}) = \mathbf{S}f(\mathbf{\Lambda})\mathbf{S}^{-1}$ is computed by applying f to the eigenvalues—a fact we exploit for the matrix exponential below.

12.3 Orthogonal, unitary, and Hermitian operators

Operators that interact well with an inner product are the ones that matter for physics. A *unitary* operator preserves the inner product, $\langle \mathbf{U}\mathbf{u}, \mathbf{U}\mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle$, which is equivalent to $\mathbf{U}^\dagger\mathbf{U} = \mathbf{I}$ (here $\mathbf{U}^\dagger = \overline{\mathbf{U}}^\mathsf{T}$ is the conjugate transpose); in the real case the same condition reads $\mathbf{Q}^\mathsf{T}\mathbf{Q} = \mathbf{I}$ and $\mathbf{Q}$ is *orthogonal*. Such maps preserve all lengths and angles and satisfy $|\det| = 1$.

An operator is *Hermitian* (self-adjoint) if $\mathbf{A}^\dagger = \mathbf{A}$, or in the real case *symmetric* if $\mathbf{A}^\mathsf{T} = \mathbf{A}$; equivalently $\langle \mathbf{u}, \mathbf{A}\mathbf{v}\rangle = \langle \mathbf{A}\mathbf{u}, \mathbf{v}\rangle$ for all $\mathbf{u}, \mathbf{v}$. Hermitian operators are pervasive: the moment-of-inertia tensor of a rigid body is symmetric, and in quantum mechanics every observable is represented by a Hermitian operator whose eigenvalues are the possible measured values.

12.4 The spectral theorem

Theorem 12.2 (Spectral theorem)**.** *A Hermitian operator has only real eigenvalues, eigenvectors belonging to distinct eigenvalues are orthogonal, and there exists an orthonormal basis of eigenvectors. In that basis* $\mathbf{A}$ *is diagonal, and*

$$\mathbf{A} = \sum_i \lambda_i \, |i\rangle\langle i|, \tag{12.1}$$

the spectral decomposition *as a sum of orthogonal projections weighted by the eigenvalues.*

Proof of the first two assertions. If $\mathbf{A}|v\rangle = \lambda|v\rangle$ with $|v\rangle \neq 0$, then $\lambda\langle v|v\rangle = \langle v|\mathbf{A}v\rangle = \langle \mathbf{A}v|v\rangle = \overline{\lambda}\langle v|v\rangle$, where the middle step uses self-adjointness; since $\langle v|v\rangle > 0$, we get $\lambda = \overline{\lambda}$, so λ is real. If also $\mathbf{A}|u\rangle = \mu|u\rangle$ with $\mu \neq \lambda$ (both real), then $\lambda\langle u|v\rangle = \langle u|\mathbf{A}v\rangle = \langle \mathbf{A}u|v\rangle = \mu\langle u|v\rangle$, so $(\lambda - \mu)\langle u|v\rangle = 0$ forces $\langle u|v\rangle = 0$. □

The spectral theorem is the finite-dimensional ancestor of the eigenfunction-expansion methods: a symmetric operator is *diagonalized by an orthonormal basis*, exactly as the operator ∂_x^2 was diagonalized by the Fourier basis. Expanding a vector in the eigenbasis and applying (12.1) turns the action of $\mathbf{A}$ into simple multiplication, eigenvalue by eigenvalue.

Example 12.3. The symmetric matrix $\mathbf{A} = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$ has characteristic equation $(2-\lambda)^2 - 1 = 0$, eigenvalues $\lambda = 3$ and $\lambda = 1$, with orthogonal eigenvectors $\frac{1}{\sqrt{2}}(1,1)$ and $\frac{1}{\sqrt{2}}(1,-1)$. Thus $\mathbf{A} = 3|e_1\rangle\langle e_1| + 1|e_2\rangle\langle e_2|$, and the orthonormal eigenvectors assemble into an orthogonal $\mathbf{S}$ whose inverse is simply its transpose.

12.5 The matrix exponential

The series that defines the ordinary exponential applies verbatim to a square matrix:

$$\exp(\mathbf{A}) = \sum_{k=0}^{\infty} \frac{\mathbf{A}^k}{k!},$$

a series that converges for every $\mathbf{A}$. It satisfies $\exp(\mathbf{A})\exp(\mathbf{B}) = \exp(\mathbf{A}+\mathbf{B})$ *when* $\mathbf{A}$ *and* $\mathbf{B}$ *commute* (and generally not otherwise). For a diagonalizable $\mathbf{A} = \mathbf{S}\boldsymbol{\Lambda}\mathbf{S}^{-1}$ the exponential is computed eigenvalue by eigenvalue,

$$\exp(t\mathbf{A}) = \mathbf{S}\,\operatorname{diag}\!\big(\mathrm{e}^{\lambda_1 t}, \ldots, \mathrm{e}^{\lambda_n t}\big)\,\mathbf{S}^{-1},$$

and for a non-diagonalizable block such as $\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)$ one finds $\exp\!\big(t\left(\begin{smallmatrix} \lambda & 1 \\ 0 & \lambda \end{smallmatrix}\right)\big) = \mathrm{e}^{\lambda t}\left(\begin{smallmatrix} 1 & t \\ 0 & 1 \end{smallmatrix}\right)$.

12.6 Linear systems of differential equations

The matrix exponential exists precisely to solve linear systems with constant coefficients. The initial value problem

$$\mathbf{x}'(t) = \mathbf{A}\,\mathbf{x}(t), \qquad \mathbf{x}(0) = \mathbf{x}_0,$$

has the unique solution $\mathbf{x}(t) = \exp(t\mathbf{A})\,\mathbf{x}_0$, as one verifies by differentiating the series term by term. When $\mathbf{A}$ is diagonalizable with eigenpairs $(\lambda_i, \mathbf{v}_i)$ this reads

$$\mathbf{x}(t) = \sum_i c_i\, \mathrm{e}^{\lambda_i t}\, \mathbf{v}_i,$$

the constants c_i fixed by the initial condition—the system's *normal modes*, each evolving independently at its own rate.

Example 12.4. For $\mathbf{A} = \left(\begin{smallmatrix} 1 & 1 \\ 4 & 1 \end{smallmatrix}\right)$ the characteristic equation $(1-\lambda)^2 - 4 = 0$ gives $\lambda = 3$ (eigenvector $(1,2)$) and $\lambda = -1$ (eigenvector $(1,-2)$), so the general solution of $\mathbf{x}' = \mathbf{A}\mathbf{x}$ is

$$\mathbf{x}(t) = c_1\, \mathrm{e}^{3t} \begin{pmatrix} 1 \\ 2 \end{pmatrix} + c_2\, \mathrm{e}^{-t} \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

The growing mode e^{3t} and the decaying mode e^{-t} are the two independent behaviors superposed in every solution.

Exercises

Exercise 12.1. Find the eigenvalues and eigenvectors of $\left(\begin{smallmatrix} 0 & 1 \\ -2 & -3 \end{smallmatrix}\right)$, and use them to write $\mathbf{A} = \mathbf{S}\mathbf{\Lambda}\mathbf{S}^{-1}$ explicitly.

Exercise 12.2. Show that the eigenvalues of a unitary operator all have modulus 1.

Exercise 12.3. Diagonalize $\mathbf{A} = \left(\begin{smallmatrix} 4 & 1 \\ 1 & 4 \end{smallmatrix}\right)$ and compute $\exp(t\mathbf{A})$.

Exercise 12.4. Solve the initial value problem $\mathbf{x}' = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)\mathbf{x}$ with $\mathbf{x}(0) = (1,0)$, and recognize the solution as a rotation.

Exercise 12.5. Verify directly that $\dfrac{d}{dt}\exp(t\mathbf{A}) = \mathbf{A}\exp(t\mathbf{A})$ by differentiating the defining series term by term.

Exercise 12.6. *Show that for any square matrix, $\det\exp(\mathbf{A}) = \mathrm{e}^{\operatorname{tr}\mathbf{A}}$. (Use diagonalization, or the fact that both sides are continuous and the identity holds on the dense set of diagonalizable matrices.)

Chapter 13

Curvilinear Coordinates and the Metric Tensor

13.1 Plane polar coordinates

A point in the plane with polar coordinates (r, θ) has position $x = r\cos\theta$, $y = r\sin\theta$. The orthonormal basis adapted to these coordinates points in the directions of increasing r and increasing θ:

$$\mathbf{e}_r = \cos\theta\,\mathbf{i} + \sin\theta\,\mathbf{j}, \qquad \mathbf{e}_\theta = -\sin\theta\,\mathbf{i} + \cos\theta\,\mathbf{j}.$$

Unlike the fixed Cartesian basis, these vectors rotate with the angle:

$$\frac{d\mathbf{e}_r}{d\theta} = \mathbf{e}_\theta, \qquad \frac{d\mathbf{e}_\theta}{d\theta} = -\mathbf{e}_r.$$

The position vector is $\mathbf{r} = r\,\mathbf{e}_r$. Differentiating with respect to time, with r and θ functions of t, gives the velocity and acceleration

$$\mathbf{v} = \dot{r}\,\mathbf{e}_r + r\dot{\theta}\,\mathbf{e}_\theta, \qquad \mathbf{a} = (\ddot{r} - r\dot{\theta}^2)\,\mathbf{e}_r + (r\ddot{\theta} + 2\dot{r}\dot{\theta})\,\mathbf{e}_\theta.$$

The term $-r\dot{\theta}^2$ is the centripetal acceleration and $2\dot{r}\dot{\theta}$ the Coriolis term. For a central force, directed along $\mathbf{e}_r$, the $\mathbf{e}_\theta$ component of $\mathbf{a}$ vanishes; since $r\ddot{\theta} + 2\dot{r}\dot{\theta} = \frac{1}{r}\frac{d}{dt}(r^2\dot{\theta})$, this gives $r^2\dot{\theta} = \text{const}$, the conservation of angular momentum (Kepler's second law).

13.2 Scale factors and the line element

For a general orthogonal coordinate system (u^1, u^2, u^3) with position $\mathbf{r}(u^1, u^2, u^3)$, the coordinate basis vectors $\partial\mathbf{r}/\partial u^i$ are mutually orthogonal. Writing

$$\frac{\partial \mathbf{r}}{\partial u^i} = h_i\,\mathbf{e}_i, \qquad h_i = \left|\frac{\partial \mathbf{r}}{\partial u^i}\right|,$$

with $\{\mathbf{e}_i\}$ orthonormal, defines the *scale factors* h_i. An infinitesimal displacement is $d\mathbf{r} = \sum_i h_i\,\mathbf{e}_i\,du^i$, so the *line element* is

$$ds^2 = d\mathbf{r}\cdot d\mathbf{r} = \sum_i h_i^2\,(du^i)^2,$$

and the volume element is $dV = h_1 h_2 h_3\,du^1\,du^2\,du^3$. The three standard systems have scale factors

$$\text{polar/cylindrical: } (1,\, r,\, 1), \qquad \text{spherical: } h_r = 1,\ h_\theta = r,\ h_\varphi = r\sin\theta,$$

giving the spherical line element $ds^2 = dr^2 + r^2\,d\theta^2 + r^2\sin^2\theta\,d\varphi^2$ and volume element $dV = r^2\sin\theta\,dr\,d\theta\,d\varphi$.

13.3 Differential operators in orthogonal coordinates

The scale factors determine the vector operators. The gradient, divergence, and Laplacian in orthogonal curvilinear coordinates are

$$\nabla\phi = \sum_i \frac{1}{h_i}\frac{\partial\phi}{\partial u^i}\,\mathbf{e}_i, \qquad \nabla\cdot\mathbf{A} = \frac{1}{h_1h_2h_3}\left[\frac{\partial(h_2h_3A_1)}{\partial u^1} + \frac{\partial(h_1h_3A_2)}{\partial u^2} + \frac{\partial(h_1h_2A_3)}{\partial u^3}\right],$$

$$\nabla^2\phi = \frac{1}{h_1h_2h_3}\sum_i \frac{\partial}{\partial u^i}\left(\frac{h_1h_2h_3}{h_i^2}\frac{\partial\phi}{\partial u^i}\right).$$

In spherical coordinates the Laplacian becomes

$$\nabla^2\phi = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\phi}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\phi}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2\phi}{\partial\varphi^2},$$

the form in which Laplace's and the wave equations are separated for problems with spherical symmetry.

13.4 General coordinates and the metric tensor

When the coordinates are not orthogonal, the coordinate basis vectors $\mathbf{g}_i = \partial\mathbf{r}/\partial u^i$ are neither perpendicular nor of unit length, and the geometry is encoded in the *metric tensor*

$$g_{ij} = \mathbf{g}_i\cdot\mathbf{g}_j = \frac{\partial\mathbf{r}}{\partial u^i}\cdot\frac{\partial\mathbf{r}}{\partial u^j}. \tag{13.1}$$

With the summation convention (repeated indices summed), the line element is

$$ds^2 = g_{ij}\,du^i\,du^j,$$

and the volume (or area) element is $dV = \sqrt{\det(g_{ij})}\,du^1\cdots du^n$. For orthogonal coordinates g_{ij} is diagonal with $g_{ii} = h_i^2$; in polar coordinates, for instance, $g = \mathrm{diag}(1, r^2)$, so $\det g = r^2$ and $dA = r\,dr\,d\theta$.

13.5 The metric as a tensor

Under a change of coordinates $u^i \to u'^a$, the chain rule $\partial\mathbf{r}/\partial u'^a = (\partial u^i/\partial u'^a)\,\partial\mathbf{r}/\partial u^i$ gives the transformation law of the metric components,

$$g'_{ab} = \frac{\partial u^i}{\partial u'^a}\frac{\partial u^j}{\partial u'^b}\,g_{ij}.$$

A quantity carrying a factor $\partial u/\partial u'$ for each lower index in this way is a *covariant tensor* of rank two; the metric is its prototype. The matrix inverse g^{ij} is a contravariant rank-two tensor, and the two are used to lower and raise indices,

$$v_i = g_{ij}v^j, \qquad v^i = g^{ij}v_j,$$

converting between the components of a vector along the coordinate directions and those along the dual directions.

Exercises

Exercise 13.1. Starting from $\mathbf{r} = r\mathbf{e}_r$ and the derivatives of $\mathbf{e}_r, \mathbf{e}_\theta$, derive the polar expressions for $\mathbf{v}$ and $\mathbf{a}$ in full.

Exercise 13.2. For parabolic coordinates $x = \frac{1}{2}(\xi^2 - \eta^2)$, $y = \xi\eta$, compute the scale factors h_ξ, h_η and the line element ds^2.

Exercise 13.3. Write out the Laplacian $\nabla^2\phi$ in cylindrical coordinates (ρ, φ, z) from the orthogonal-coordinate formula.

Exercise 13.4. For spherical coordinates, compute the metric tensor g_{ij} from (13.1) and confirm $\sqrt{\det g} = r^2 \sin\theta$.

Exercise 13.5. Verify the transformation law $g'_{ab} = (\partial u^i/\partial u'^a)(\partial u^j/\partial u'^b)g_{ij}$ for the change from Cartesian to polar coordinates.

Exercise 13.6. *Show that for any coordinate system $g^{ik}g_{kj} = \delta^i_j$, and use this to verify that raising then lowering an index returns the original components.

Chapter 14

The Calculus of Variations

14.1 Functionals and the variational problem

Many laws of physics and geometry single out a curve or function that extremizes an integral. Fermat's principle states that light follows the path of least travel time; the brachistochrone problem asks for the curve of fastest descent of a bead under gravity; a geodesic is the shortest path between two points. Each seeks the function $y(x)$ that extremizes a *functional*

$$I[y] = \int_a^b F(x, y, y')\, dx, \tag{14.1}$$

subject to fixed endpoint values $y(a) = y_a$ and $y(b) = y_b$. A functional assigns a number to each admissible function, and the calculus of variations locates its stationary points.

14.2 The Euler–Lagrange equation

Let y extremize I, and compare it with the nearby functions $y(x) + \varepsilon\eta(x)$, where $\eta(a) = \eta(b) = 0$ so that the endpoints are held fixed. Then $I[y + \varepsilon\eta]$ is an ordinary function of ε with a stationary point at $\varepsilon = 0$:

$$0 = \frac{d}{d\varepsilon} I[y + \varepsilon\eta]\Big|_{\varepsilon=0} = \int_a^b \big(F_y\, \eta + F_{y'}\, \eta'\big)\, dx.$$

Integrating the second term by parts and using $\eta(a) = \eta(b) = 0$ removes the boundary contribution, leaving

$$\int_a^b \left(F_y - \frac{d}{dx} F_{y'}\right) \eta\, dx = 0.$$

Because η is arbitrary, the bracket must vanish identically. This is the *Euler–Lagrange equation*

$$F_y - \frac{d}{dx} F_{y'} = 0, \tag{14.2}$$

a second-order differential equation for the extremizing function $y(x)$.

14.3 First integrals

Two situations reduce the order of (14.2). If F does not contain y, then $F_y = 0$ and the equation integrates immediately to $F_{y'} = \text{const}$. If F does not contain x explicitly, then the *Beltrami identity*

$$F - y' F_{y'} = \text{const}$$

holds, as one checks by differentiating the left side and using (14.2).

Example 14.1 (Shortest path)**.** The arc length functional has $F = \sqrt{1+y'^2}$, independent of y, so $F_{y'} = y'/\sqrt{1+y'^2}$ is constant; hence y' is constant and the extremal is a straight line.

Example 14.2 (Minimal surface of revolution)**.** Minimizing the area $2\pi \int y\sqrt{1+y'^2}\,dx$ uses $F = y\sqrt{1+y'^2}$, which has no explicit x. The Beltrami identity gives $y/\sqrt{1+y'^2} = c$, and separating variables integrates to $y = c\cosh\big((x-x_0)/c\big)$: the minimal surface is the catenoid generated by a catenary.

Example 14.3 (Brachistochrone)**.** The descent time is $\int \sqrt{(1+y'^2)/(2gy)}\,dx$, again with no explicit x. The Beltrami identity reduces to $y(1+y'^2) = \text{const}$, whose solution is a cycloid $x = a(\varphi - \sin\varphi)$, $y = a(1-\cos\varphi)$—the curve of fastest descent.

14.4 Several variables and Lagrangian mechanics

When F depends on several functions $y_1, \ldots, y_n$ and their derivatives, the variational argument applied to each independently yields one Euler–Lagrange equation per function,

$$\frac{\partial F}{\partial y_k} - \frac{d}{dx}\frac{\partial F}{\partial y'_k} = 0, \qquad k = 1, \ldots, n.$$

In mechanics the independent variable is time and the functional is the *action*

$$S[q] = \int L(q, \dot q, t)\,dt, \qquad L = T - U,$$

the Lagrangian being the kinetic minus the potential energy. Hamilton's principle states that the physical trajectory extremizes S, so the equations of motion are the Euler–Lagrange equations

$$\frac{d}{dt}\frac{\partial L}{\partial \dot q_k} - \frac{\partial L}{\partial q_k} = 0,$$

with $p_k = \partial L/\partial \dot q_k$ the generalized momentum.

Example 14.4. For the harmonic oscillator $L = \frac{1}{2}m\dot x^2 - \frac{1}{2}kx^2$, the Euler–Lagrange equation gives $m\ddot x = -kx$. For the simple pendulum $L = \frac{1}{2}m\ell^2\dot\theta^2 + mg\ell\cos\theta$, it gives $\ell\ddot\theta = -g\sin\theta$. A double pendulum is described by two generalized coordinates and yields two coupled equations.

14.5 Fields: the vibrating membrane

For a quantity depending on several independent variables—a field—the functional is an integral over a region and the Euler–Lagrange equation is a partial differential equation. The transverse displacement $u(x, y, t)$ of a membrane under tension τ and with area density ρ has the action

$$S = \iiint \left[\tfrac{1}{2}\rho\, u_t^2 - \tfrac{1}{2}\tau\big(u_x^2 + u_y^2\big)\right] dA\,dt,$$

with Lagrangian density $\mathcal{L}$. The field Euler–Lagrange equation

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial u_t} + \frac{\partial}{\partial x}\frac{\partial \mathcal{L}}{\partial u_x} + \frac{\partial}{\partial y}\frac{\partial \mathcal{L}}{\partial u_y} - \frac{\partial \mathcal{L}}{\partial u} = 0$$

applied to this density gives $\rho\, u_{tt} = \tau(u_{xx} + u_{yy})$, that is the wave equation $u_{tt} = c^2\nabla^2 u$ with $c^2 = \tau/\rho$. The wave equation of Chapter 3 is thus the Euler–Lagrange equation of the membrane's action.

Exercises

Exercise 14.1. Write the Euler–Lagrange equation for $F = \frac{1}{2}y'^2 - V(y)$ and identify it as Newton's equation for a particle in the potential V.

Exercise 14.2. Show that the geodesics of the plane (extremals of arc length) are straight lines, working directly from (14.2).

Exercise 14.3. Derive the catenoid $y = c\cosh((x - x_0)/c)$ as the minimal surface of revolution, supplying the integration omitted in the text.

Exercise 14.4. From the pendulum Lagrangian $L = \frac{1}{2}m\ell^2\dot{\theta}^2 + mg\ell\cos\theta$, obtain the equation of motion and its small-angle approximation.

Exercise 14.5. *Carry out the Beltrami reduction of the brachistochrone functional to $y(1 + y'^2) =$ const, and verify that the cycloid satisfies it.

Exercise 14.6. *Apply the field Euler–Lagrange equation to the membrane Lagrangian density and obtain the wave equation, identifying the wave speed.

Chapter 15

Constrained Variation and Lagrange Multipliers

15.1 Lagrange multipliers in finite dimensions

Consider extremizing a function $f(\mathbf{x})$ of $\mathbf{x} \in \mathbb{R}^n$ subject to a constraint $g(\mathbf{x}) = 0$. At a constrained extremum the gradient ∇f can have no component tangent to the surface $g = 0$; otherwise a small move along the surface would change f. Since ∇g is normal to that surface, ∇f must be parallel to it, so there is a scalar λ, the *Lagrange multiplier*, with

$$\nabla f = \lambda \nabla g. \tag{15.1}$$

Equation (15.1) together with the constraint $g(\mathbf{x}) = 0$ provides $n + 1$ equations for the $n + 1$ unknowns $\mathbf{x}$ and λ.

Example 15.1. To extremize $f = x + y$ on the circle $x^2 + y^2 = 1$, (15.1) gives $(1, 1) = \lambda(2x, 2y)$, so $x = y$; the constraint then yields $x = y = \pm 1/\sqrt{2}$ and the extreme values $f = \pm\sqrt{2}$.

15.2 Isoperimetric constraints

The same device handles a functional extremized under an integral constraint. To extremize $I[y] = \int_a^b F\,dx$ subject to $J[y] = \int_a^b G\,dx = \text{const}$, introduce a multiplier λ and extremize the combined functional $\int_a^b (F - \lambda G)\,dx$ without constraint. Its Euler–Lagrange equation is

$$\frac{\partial(F - \lambda G)}{\partial y} - \frac{d}{dx}\frac{\partial(F - \lambda G)}{\partial y'} = 0,$$

with λ determined afterward by the constraint. The classical isoperimetric problem—maximizing the enclosed area for a fixed perimeter—is of this type, and its solution is the circle.

15.3 The catenary

A uniform flexible chain of linear density ρ and fixed length L hangs between two points and settles into the shape that minimizes its gravitational potential energy

$$U = \rho g \int y\sqrt{1 + y'^2}\,dx$$

subject to the length constraint $\int \sqrt{1+y'^2}\,dx = L$. With a multiplier λ, the combined integrand $(\rho g\,y - \lambda)\sqrt{1+y'^2}$ has no explicit x, so the Beltrami identity applies and gives

$$\frac{\rho g\,y - \lambda}{\sqrt{1+y'^2}} = \text{const.}$$

Integrating yields $y = \dfrac{\lambda}{\rho g} + c\cosh\dfrac{x - x_0}{c}$: the hanging chain is a catenary, a shifted hyperbolic cosine.

15.4 Constraints in mechanics

Holonomic constraints in mechanics are imposed in the same way. A constraint $g(\mathbf{q}) = 0$ on the generalized coordinates is enforced by adding $\lambda(t)\,g(\mathbf{q})$ to the Lagrangian and varying freely; the resulting multiplier $\lambda(t)$ is proportional to the constraint force that the constraint surface exerts to keep the motion on it.

Exercises

Exercise 15.1. Find the maximum of $f = xyz$ subject to $x + y + z = 1$ with $x, y, z > 0$ using Lagrange multipliers.

Exercise 15.2. Find the rectangle of largest area inscribed with sides parallel to the axes in the ellipse $x^2/a^2 + y^2/b^2 = 1$.

Exercise 15.3. Show, using the isoperimetric method, that among closed curves of fixed length the circle encloses the greatest area (set up the multiplier problem and identify the extremal).

Exercise 15.4. Complete the integration of the catenary equation $(\rho g\,y - \lambda)/\sqrt{1+y'^2} = \text{const}$ to obtain the hyperbolic cosine.

Exercise 15.5. *A bead slides on a wire constrained to the curve $g(x, y) = 0$ under gravity. Set up the Lagrangian with a multiplier $\lambda(t)$ and interpret λ as the normal force exerted by the wire.

Solutions to Selected Exercises

Chapter 1.

1.1. If f is even and g is odd then $(fg)(-x) = f(-x)g(-x) = f(x)\big(-g(x)\big) = -(fg)(x)$, so fg is odd; the product of two odd functions is even by the same computation with two sign changes. For an odd f, each integrand $f(x)\cos(n\pi x/L)$ is odd, so it integrates to zero over $[-L, L]$ and $a_n = 0$.

1.2. Oddness gives $a_n = 0$ and $b_n = \frac{2}{L}\int_0^L x\sin(n\pi x/L)\,dx$. Integration by parts gives the value $(-1)^{n+1}L^2/(n\pi)$ for that integral, hence $b_n = (-1)^{n+1}\,2L/(n\pi)$, the stated series. The coefficients decay only like $1/n$ because the *periodic extension* has a jump of size $2L$ at the odd multiples of L. Smoothness on the open interval is irrelevant; what governs the decay rate is the regularity of the periodic extension.

1.4. One finds $a_0/2 = L^2/3$ and, by parts twice, $a_n = 4L^2(-1)^n/(n^2\pi^2)$, so

$$x^2 = \frac{L^2}{3} + \frac{4L^2}{\pi^2}\sum_{n\geq 1}\frac{(-1)^n}{n^2}\cos\frac{n\pi x}{L}.$$

Evaluating at $x = L$ (a point of continuity of the even extension), where $\cos(n\pi) = (-1)^n$, gives $L^2 = \frac{L^2}{3} + \frac{4L^2}{\pi^2}\sum_{n\geq 1}\frac{1}{n^2}$, and solving yields $\sum_{n\geq 1} n^{-2} = \pi^2/6$.

1.5. For the square wave, $\frac{1}{2L}\int_{-L}^{L} 1\,dx = 1$, while the right side of (1.8) is $\frac{1}{2}\sum_n b_n^2 = \frac{1}{2}\sum_{k\geq 0}(4/((2k+1)\pi))^2 = \frac{8}{\pi^2}\sum_{k\geq 0}(2k+1)^{-2}$. Setting the two equal gives $\sum_{k\geq 0}(2k+1)^{-2} = \pi^2/8$.

Chapter 2.

2.3. With α general, $\widetilde{f}(k) = \sqrt{\pi/\alpha}\,\mathrm{e}^{-k^2/4\alpha}$. Then $|f|^2 = \mathrm{e}^{-2\alpha x^2}$ has $(\Delta x)^2 = 1/4\alpha$, while $\left|\widetilde{f}\right|^2 \propto \mathrm{e}^{-k^2/2\alpha}$ has $(\Delta k)^2 = \alpha$; the product is $(\Delta x)^2(\Delta k)^2 = 1/4$, so $\Delta x\,\Delta k = 1/2$, the minimum allowed by Theorem 2.8.

2.5. The integrand of $(f * f)(x) = \int f(y)f(x-y)\,dy$ is 1 exactly where both $|y| < a$ and $|x-y| < a$, i.e. on an interval whose length is the overlap of $[-a, a]$ with its translate by x. That length is $2a - |x|$ for $|x| < 2a$ and zero otherwise, giving the triangle.

2.6. Transforming $-u'' + a^2u = f$ and using $\mathcal{F}[u''] = -k^2\widetilde{u}$ gives $(k^2 + a^2)\widetilde{u} = \widetilde{f}$, so $\widetilde{u} = \widetilde{f}/(k^2 + a^2)$. Since $\mathcal{F}[\mathrm{e}^{-a|x|}] = 2a/(k^2 + a^2)$ (Exercise 2.1), the factor $1/(k^2 + a^2)$ is the transform of $G(x) = \mathrm{e}^{-a|x|}/(2a)$. By the convolution theorem $u = G * f$, that is $u(x) = \frac{1}{2a}\int_{-\infty}^{\infty}\mathrm{e}^{-a|x-y|}f(y)\,dy$.

2.7. For the box of half-width $a = 1$, $\int |f|^2\,dx = 2$ and $\widetilde{f}(k) = 2\sin k/k$. Plancherel gives $2 = \frac{1}{2\pi}\int(2\sin k/k)^2\,dk = \frac{2}{\pi}\int(\sin k/k)^2\,dk$, hence $\int_{-\infty}^{\infty}(\sin k/k)^2\,dk = \pi$.

Chapter 3.

3.1. Here $\widetilde{f}(k) = \sqrt{\pi}\,\mathrm{e}^{-k^2/4}$, so $\widetilde{u}(k,t) = \sqrt{\pi}\,\mathrm{e}^{-k^2(1+4\kappa t)/4}$, the transform of a Gaussian. Inverting, $u(x,t) = (1+4\kappa t)^{-1/2}\mathrm{e}^{-x^2/(1+4\kappa t)}$: the width grows like $\sqrt{1+4\kappa t}$ and the height falls like $(1+4\kappa t)^{-1/2}$, consistent with the diffusive spreading of (3.2).

3.3. Write $\Phi = (4\pi\kappa t)^{-1/2}\mathrm{e}^{-x^2/4\kappa t}$. Direct differentiation gives $\Phi_t = \Phi\big(-\frac{1}{2t} + \frac{x^2}{4\kappa t^2}\big)$ and $\Phi_{xx} = \Phi\big(-\frac{1}{2\kappa t} + \frac{x^2}{4\kappa^2 t^2}\big)$, and one checks $\Phi_t = \kappa\Phi_{xx}$. The normalization $\int \Phi\,dx = 1$ is the standard Gaussian integral $\int \mathrm{e}^{-x^2/4\kappa t}\,dx = \sqrt{4\pi\kappa t}$.

3.4. By Exercise 2.1, $\mathrm{e}^{-a|x|} = \frac{1}{2\pi}\int \frac{2a}{k^2+a^2}\mathrm{e}^{\mathrm{i}kx}\,dk$. Renaming variables ($x \leftrightarrow k$, $a \to y$) and rearranging shows $\int (1/\pi)\,y/(x^2+y^2)\,\mathrm{e}^{-\mathrm{i}kx}\,dx = \mathrm{e}^{-|k|y}$, which is the claim.

Chapter 4.

4.2. Write $-16 = 16\,\mathrm{e}^{\mathrm{i}\pi}$; the fourth roots are $2\,\mathrm{e}^{\mathrm{i}(\pi+2\pi k)/4}$ for $k = 0,1,2,3$, namely

$$2\mathrm{e}^{\mathrm{i}\pi/4},\quad 2\mathrm{e}^{\mathrm{i}3\pi/4},\quad 2\mathrm{e}^{\mathrm{i}5\pi/4},\quad 2\mathrm{e}^{\mathrm{i}7\pi/4},\qquad \text{i.e.}\qquad \sqrt{2}(\pm 1 \pm \mathrm{i}).$$

They lie on the circle of radius 2 at the corners of a square.

4.4. $\operatorname{Log}(-\mathrm{i}) = \ln 1 + \mathrm{i}\operatorname{Arg}(-\mathrm{i}) = -\mathrm{i}\pi/2$. For $(1+\mathrm{i})^{\mathrm{i}}$, use $\log(1+\mathrm{i}) = \frac{1}{2}\ln 2 + \mathrm{i}(\pi/4 + 2\pi k)$; then $(1+\mathrm{i})^{\mathrm{i}} = \mathrm{e}^{\mathrm{i}\log(1+\mathrm{i})} = \mathrm{e}^{-(\pi/4+2\pi k)}\mathrm{e}^{\mathrm{i}\frac{1}{2}\ln 2}$, with principal value $\mathrm{e}^{-\pi/4}(\cos(\frac{1}{2}\ln 2) + \mathrm{i}\sin(\frac{1}{2}\ln 2))$.

4.5. From $\mathrm{e}^{\mathrm{i}z} = \mathrm{e}^{-y}(\cos x + \mathrm{i}\sin x)$ one finds $\sin z = \sin x\cosh y + \mathrm{i}\cos x\sinh y$, so $|\sin z|^2 = \sin^2 x\cosh^2 y + \cos^2 x\sinh^2 y = \sin^2 x + \sinh^2 y$ (using $\cosh^2 - \sinh^2 = 1$). As $|y| \to \infty$ this grows without bound, so $\sin z$ is unbounded on $\mathbb{C}$.

Chapter 5.

5.2. $u = \mathrm{e}^x\cos y$ gives $u_{xx} = \mathrm{e}^x\cos y$, $u_{yy} = -\mathrm{e}^x\cos y$, so $u_{xx} + u_{yy} = 0$. From $v_y = u_x = \mathrm{e}^x\cos y$ we get $v = \mathrm{e}^x\sin y + g(x)$; then $v_x = \mathrm{e}^x\sin y + g'(x) = -u_y = \mathrm{e}^x\sin y$ forces $g' \equiv 0$. Thus $v = \mathrm{e}^x\sin y$ and $f = \mathrm{e}^x(\cos y + \mathrm{i}\sin y) = \mathrm{e}^z$.

5.3. If $f' \equiv 0$ then $u_x = v_x = 0$ and, by (5.1), $u_y = v_y = 0$ as well; all partials vanish on the (connected) domain, so u and v are constant, hence so is f.

5.4. Differentiating $u^2 + v^2 = c$ gives $uu_x + vv_x = 0$ and $uu_y + vv_y = 0$. Using (5.1) to replace u_y, v_y turns this into a homogeneous linear system for (u_x, v_x) with determinant $u^2 + v^2$. If $c \neq 0$ the only solution is $u_x = v_x = 0$, so $f' = 0$ and f is constant (the case $c = 0$ gives $f \equiv 0$).

Chapter 6.

6.1. The point $z = 1$ lies inside $|z| = 2$, so $\oint dz/(z-1) = 2\pi\mathrm{i}$ by the deformation principle; $z = 3$ lies outside, the integrand is analytic inside the contour, and Cauchy's theorem gives 0.

6.3. By Cauchy's integral formula with $f(z) = \mathrm{e}^z$ and $a = 0$, $\oint_{|z|=1} \mathrm{e}^z/z\,dz = 2\pi\mathrm{i}\,\mathrm{e}^0 = 2\pi\mathrm{i}$; with $f(z) = \cos z$ and $a = \mathrm{i}$, $\oint_{|z|=2} \cos z/(z-\mathrm{i})\,dz = 2\pi\mathrm{i}\cos\mathrm{i} = 2\pi\mathrm{i}\cosh 1$.

6.4. By (6.4) with $n = 2$ and $f(z) = \mathrm{e}^z$, $\oint_{|z|=1} \mathrm{e}^z/z^3\,dz = \dfrac{2\pi\mathrm{i}}{2!}f''(0) = \pi\mathrm{i}$.

Chapter 7.

7.2. Partial fractions: $\dfrac{1}{z(z-1)} = \dfrac{1}{z-1} - \dfrac{1}{z}$. (a) For $0 < |z| < 1$, expand $\dfrac{1}{z-1} = -\dfrac{1}{1-z} = -\sum_{n\geq 0} z^n$, so $f(z) = -\dfrac{1}{z} - \sum_{n\geq 0} z^n$ (a simple pole at 0, residue -1). (b) For $|z| > 1$, expand $\dfrac{1}{z-1} = \dfrac{1}{z}\dfrac{1}{1-1/z} = \sum_{n\geq 1} z^{-n}$, so $f(z) = \sum_{n\geq 1} z^{-n} - \dfrac{1}{z} = \sum_{n\geq 2} z^{-n}$.

7.3. $(1-\cos z)/z^2 = \frac{1}{2} - \frac{z^2}{24} + \cdots$: removable, residue 0. $\sin z/z^3 = z^{-2} - \frac{1}{6} + \cdots$: pole of order 2, residue 0. $z\,\mathrm{e}^{1/z} = z\sum_{n\geq 0} z^{-n}/n! = z + 1 + \frac{1}{2z} + \cdots$: essential singularity, residue $\frac{1}{2}$.

7.5. For $\oint_{|z|=3} dz/(z^2+1)$, both poles $\pm\mathrm{i}$ lie inside; residues $1/(2\mathrm{i})$ and $-1/(2\mathrm{i})$ sum to 0, so the integral is 0. For the second, the poles inside are $z=1$ (order 2) and $z=\mathrm{i}$ (simple); summing $2\pi\mathrm{i}$ times their residues gives the value.

Chapter 8.

8.1. The integrand $1/(z^2+1)^2$ has a double pole at $z=\mathrm{i}$ in the upper half-plane. By Proposition 7.6,

$$\mathrm{Res}_{z=\mathrm{i}} = \frac{d}{dz}\frac{1}{(z+\mathrm{i})^2}\bigg|_{z=\mathrm{i}} = -\frac{2}{(2\mathrm{i})^3} = \frac{1}{4\mathrm{i}},$$

so the integral is $2\pi\mathrm{i}\cdot\dfrac{1}{4\mathrm{i}} = \dfrac{\pi}{2}$.

8.3. $\int \mathrm{e}^{\mathrm{i}x}/(x^2+4)\,dx = 2\pi\mathrm{i}\ \mathrm{Res}_{z=2\mathrm{i}} = 2\pi\mathrm{i}\cdot\dfrac{\mathrm{e}^{-2}}{4\mathrm{i}} = \dfrac{\pi}{2}\mathrm{e}^{-2}$; the cosine integral is its real part, $\frac{\pi}{2}\mathrm{e}^{-2}$.

8.4. With $a=5, b=4$, the inside root is $z_+ = (-5+\sqrt{25-16})/4 = -\frac{1}{2}$, and the formula $2\pi/\sqrt{a^2-b^2} = 2\pi/3$ gives the value $\frac{2\pi}{3}$.

Chapter 9.

9.2. Equate cross-ratios, $(z,0,1,\infty) = (w,1,\mathrm{i},-1)$. Using $(z,0,1,\infty) = z$ and $(w,1,\mathrm{i},-1) = \dfrac{(w-1)(\mathrm{i}+1)}{(w+1)(\mathrm{i}-1)}$, solve $\dfrac{(w-1)(\mathrm{i}+1)}{(w+1)(\mathrm{i}-1)} = z$ for w to obtain the required map. Since a Mobius transformation carries circles-and-lines to circles-and-lines and the images $1, \mathrm{i}, -1$ all lie on the unit circle, the real axis is sent to that circle.

9.5. $\Omega'(z) = U(1-a^2/z^2)$, so the velocity vanishes where $z^2 = a^2$, i.e. at $z = \pm a$. Both lie on the cylinder $|z| = a$; these are the front and rear stagnation points of the flow.

9.6. With $f = u + \mathrm{i}v$ analytic, the chain rule gives $\Psi_x = \Phi_u u_x + \Phi_v v_x$ and a second differentiation, plus $u_{xx} + u_{yy} = v_{xx} + v_{yy} = 0$ and the Cauchy–Riemann relations $u_x = v_y$, $u_y = -v_x$, collapses the cross terms and yields $\Psi_{xx} + \Psi_{yy} = (u_x^2 + u_y^2)(\Phi_{uu} + \Phi_{vv}) = |f'|^2\,(\Phi_{uu} + \Phi_{vv})$.

Chapter 10.

10.1. A symmetric 2×2 matrix is determined by its three entries $a_{11}, a_{22}, a_{12} = a_{21}$, so the space is three-dimensional with basis $\left(\begin{smallmatrix}1&0\\0&0\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0&0\\0&1\end{smallmatrix}\right)$, $\left(\begin{smallmatrix}0&1\\1&0\end{smallmatrix}\right)$. Closure under addition and scaling is immediate since the transpose is linear.

10.2. They are independent and three in number, so they form a basis of the (three-dimensional) space. Solving $x^2 = a(1) + b(1+x) + c(1+x+x^2)$ gives $c = 1, b = -1, a = 0$, i.e. $x^2 = (1+x+x^2) - (1+x)$.

10.5. With $T = d/dx$ on $\{1, x, x^2, x^3\}$, $T1 = 0$, $Tx = 1$, $Tx^2 = 2x$, $Tx^3 = 3x^2$, so the matrix is strictly upper triangular with super-diagonal $1, 2, 3$. Hence $\det T = 0$ (it is not invertible—constants are killed) and $\operatorname{tr} T = 0$.

Chapter 11.

11.2. Gram–Schmidt on $1, x, x^2$ with $w = 1$ on $[-1, 1]$: 1 is already orthogonal to x (odd × even integrates to zero); from x^2 subtract its projection onto 1, namely $\langle 1, x^2\rangle / \langle 1, 1\rangle = \frac{2/3}{2} = \frac{1}{3}$, giving $x^2 - \frac{1}{3}$. Up to normalization these are P_0, P_1, P_2.

11.3. $\int_{-L}^{L} e^{im\pi x/L}\overline{e^{in\pi x/L}}\,dx = \int_{-L}^{L} e^{i(m-n)\pi x/L}\,dx = 2L\,\delta_{mn}$, so the functions are orthogonal with squared norm $2L$; dividing by $\sqrt{2L}$ makes them orthonormal.

11.6. Apply $\|u\|^2 = \langle u|u\rangle$ and insert the identity $\sum_i |e_i\rangle\langle e_i| = \mathbf{I}$ between bra and ket: $\langle u|u\rangle = \sum_i \langle u|e_i\rangle\langle e_i|u\rangle = \sum_i |\langle e_i|u\rangle|^2$.

Chapter 12.

12.1. $\det\begin{pmatrix} -\lambda & 1 \\ -2 & -3-\lambda \end{pmatrix} = \lambda^2 + 3\lambda + 2 = (\lambda+1)(\lambda+2)$, so $\lambda = -1$ (eigenvector $(1, -1)$) and $\lambda = -2$ (eigenvector $(1, -2)$); these columns form $\mathbf{S}$ and $\mathbf{\Lambda} = \mathrm{diag}(-1, -2)$.

12.3. Eigenvalues of $\begin{pmatrix} 4 & 1 \\ 1 & 4 \end{pmatrix}$ are 5 and 3 with orthonormal eigenvectors $\frac{1}{\sqrt{2}}(1, 1), \frac{1}{\sqrt{2}}(1, -1)$, so $\exp(t\mathbf{A}) = \mathbf{S}\,\mathrm{diag}(e^{5t}, e^{3t})\mathbf{S}^{\mathsf{T}}$, which works out to $\frac{1}{2}\begin{pmatrix} e^{5t}+e^{3t} & e^{5t}-e^{3t} \\ e^{5t}-e^{3t} & e^{5t}+e^{3t} \end{pmatrix}$.

12.4. The eigenvalues of $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ are $\pm i$; summing the modes (or exponentiating directly) gives $\mathbf{x}(t) = (\cos t, -\sin t)$, the clockwise rotation of the initial vector $(1, 0)$, consistent with $\exp(t\mathbf{A}) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$.

Chapter 13.

13.3. For cylindrical coordinates $h_\rho = 1, h_\varphi = \rho, h_z = 1$, so $h_1h_2h_3 = \rho$ and the formula gives $\nabla^2\phi = \dfrac{1}{\rho}\dfrac{\partial}{\partial\rho}\Big(\rho\dfrac{\partial\phi}{\partial\rho}\Big) + \dfrac{1}{\rho^2}\dfrac{\partial^2\phi}{\partial\varphi^2} + \dfrac{\partial^2\phi}{\partial z^2}$.

13.4. With $\mathbf{r} = (r\sin\theta\cos\varphi, r\sin\theta\sin\varphi, r\cos\theta)$ the vectors $\partial\mathbf{r}/\partial r, \partial\mathbf{r}/\partial\theta, \partial\mathbf{r}/\partial\varphi$ are orthogonal with lengths $1, r, r\sin\theta$, so $g = \mathrm{diag}(1, r^2, r^2\sin^2\theta)$ and $\sqrt{\det g} = r^2\sin\theta$.

Chapter 14.

14.1. With $F = \frac{1}{2}y'^2 - V(y)$, $F_{y'} = y'$ and $F_y = -V'(y)$, so (14.2) reads $-V'(y) - \dfrac{d}{dx}y' = 0$, i.e. $y'' = -V'(y)$, which is Newton's equation $m\ddot{x} = -V'(x)$ (here $m = 1$, with x the independent variable renamed).

14.3. From $y/\sqrt{1 + y'^2} = c$, square to get $y^2 = c^2(1+y'^2)$, so $y' = \sqrt{y^2 - c^2}/c$ and $c\,dy/\sqrt{y^2 - c^2} = dx$. Integrating, $c\,\cosh^{-1}(y/c) = x - x_0$, hence $y = c\cosh\big((x - x_0)/c\big)$.

14.4. The Euler–Lagrange equation $\frac{d}{dt}(m\ell^2\dot\theta) - (-mg\ell\sin\theta) = 0$ gives $\ell\ddot\theta = -g\sin\theta$; for small θ, $\sin\theta \approx \theta$ and $\ddot\theta = -(g/\ell)\theta$, simple harmonic motion with frequency $\sqrt{g/\ell}$.

Chapter 15.

15.1. $\nabla(xyz) = \lambda\nabla(x + y + z)$ gives $yz = xz = xy = \lambda$, so $x = y = z$; the constraint $x + y + z = 1$ yields $x = y = z = \frac{1}{3}$ and maximum $f = 1/27$.

15.2. Parametrize a corner (x, y) with $x, y > 0$; the area is $4xy$ subject to $x^2/a^2 + y^2/b^2 = 1$. Then $\nabla(4xy) = \lambda\nabla(x^2/a^2 + y^2/b^2)$ gives $x = a/\sqrt{2},\ y = b/\sqrt{2}$, so the largest inscribed rectangle has area $2ab$.

15.4. Writing the constant as c, $(\rho gy - \lambda) = c\sqrt{1 + y'^2}$ gives $y' = \sqrt{(\rho gy - \lambda)^2 - c^2}/c$; separating and integrating produces $\rho g\,y - \lambda = c\cosh(\rho g(x - x_0)/c)$, i.e. the catenary up to the constant shift $\lambda/\rho g$.